The Thesis Committee for Karinne Viktoria Berstis
Certifies that this is the approved version of the following Thesis:

Investigating changes in connected speech in nonfluent/agrammatic primary progressive aphasia following script training

APPROVED BY
SUPERVISING COMMITTEE:

Maya Henry, Supervisor

Liberty Hamilton

Investigating changes in connected speech in nonfluent/agrammatic primary progressive aphasia following script training

by

Karinne Viktoria Berstis

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2020

Acknowledgements

Thank you to Maya Henry for her incredible mentorship throughout this process, and over the last few years. You continue to challenge me to expand my horizons and seek new knowledge and opportunities. I truly admire your determination in your pursuits and have enjoyed sharing in your curiosity. Thanks also to Liberty Hamilton for her valuable insights and contributions in the editing process. Your input and perspectives greatly helped to contextualize this study and I very much appreciate you taking the time to read my thesis. I would also like to thank Stephanie Grasso, who was instrumental in the completion of this project. Your patience and brilliance are unparalleled and I could not have asked for a better collaborator. Thank you to all the members of the Aphasia Research and Treatment Lab, particularly Kristin Schaffer and Willa Keegan-Rodewald for their help with transcription. Finally, thank you to Heather Dial for your emotional support and to Hayden Bagot for reminding me to eat and continually encouraging me.

Abstract

Investigating changes in connected speech in nonfluent/agrammatic primary progressive aphasia following script training

Karinne Viktoria Berstis, MA

The University of Texas at Austin, 2020

Supervisor: Maya Henry

Purpose: Script training, in which functional, scripted monologues or dialogues are trained, is an effective intervention for individuals with aphasia and apraxia of speech. It has been shown to improve production of scripted speech, allowing individuals to more effectively communicate and socially engage in conversational contexts. A previous study evaluating the efficacy of script training in individuals with nonfluent/agrammatic PPA (nfvPPA; Henry et al., 2018) documented improved production of trained scripts at post-treatment, with maintenance up to one year post-treatment. However, treatment-induced changes in discourse measures have yet to be evaluated, in part due to the time and labor-intensive nature of discourse analysis. The purpose of this study was to evaluate the utility of discourse-level measures for capturing treatment outcomes in individuals with nfvPPA who have undergone script training. Specifically, we examined speech fluency, grammar, and informativeness using metrics derived with minimal hand-coding of transcriptions.

Method: Language samples from 20 individuals (n=10 from a previous study) with nfvPPA who underwent Video-Implemented Script Training for Aphasia (VISTA) were analyzed for this study. Probes eliciting responses to trained and untrained script topics at pre- and post-treatment were transcribed and coded using Computerized Language ANalysis (CLAN, MacWhinney, 2000). Transcriptions were analyzed for words per minute (WPM), fluency disruptions per hundred words, mean length of utterance in morphemes (MLUm), grammatical complexity, propositional idea density, and proportion of open to closed class words.

Results: Participants demonstrated significant improvement for trained topics from pre- to post-treatment in WPM, fluency disruptions per hundred words, MLUm, grammatical complexity, and proportion of open to closed class words. A significant difference was also observed for untrained topics in fluency disruptions per hundred words.

Conclusion: Complementing previous findings (Henry et al., 2018), novel discourse measures revealed improvements in grammar and speech rate for trained material following VISTA. Additionally, speech fluency improved during production of both trained and untrained material. These findings lend additional support for script training as a means to improve connected speech production in individuals with nfvPPA. Further, this study illustrates the utility of discourse measures calculated automatically in characterizing treatment effects on connected speech following script training.

Table of Contents

List of Tables

List of Figures

Introduction

PRIMARY PROGRESSIVE APHASIA

Primary progressive aphasia (PPA) is a clinical syndrome characterized by progressive deterioration of speech and language with relative sparing of cognitive and non-speech motoric functions in its initial stages (Gorno-Tempini et al., 2011; Mesulam, 1982). With progression of the underlying neurodegenerative disease, individuals with PPA begin to experience more global changes, including motoric impairments, cognitive deficits, and changes in behavior or personality (Dickerson, 2011; Harciarek et al., 2014; Rogalski & Mesulam, 2009). International consensus criteria delineate three distinct PPA phenotypes, with unique clinical presentations and patterns of underlying brain atrophy (Gorno-Tempini et al., 2004, 2011; Wilson et al., 2010). Two clinical variants, semantic PPA and nonfluent/agrammatic PPA, are commonly associated with frontotemporal lobar degeneration (tauopathy or TDP-43 proteinopathy; Spinelli et al., 2017). The third variant, logopenic PPA, is most commonly associated with underlying Alzheimer's pathology (Spinelli et al., 2017).

The nonfluent/agrammatic variant of PPA (nfvPPA), which is the focus of this study, is characterized by grammatical deficits in language production and/or effortful, halting speech consistent with apraxia of speech (Ash et al., 2010; Gorno-Tempini et al., 2011; Grossman, 2012; Montembeault et al., 2018). In addition to these deficits, consensus criteria for nfvPPA dictate that two of the following characteristics must also be present: reduced comprehension of syntactically complex sentences, spared single-word comprehension, and intact object knowledge. Neuroimaging studies have shown left posterior fronto-insular atrophy to be associated with this clinical syndrome (Gorno-Tempini et al., 2011; Grossman, 2012). Speech in this population is often markedly reduced in rate and may present as telegraphic or distorted, with frequent pauses and repetitions (Ash et al., 2010, 2013; Croot et al., 2012; Santos-Santos et al., 2016; Thompson et al., 1997, 2012; Wilson et al., 2010). Deficits in grammatical ability may be noted in the spoken language of individuals with nfvPPA, with reductions in syntactic complexity and mean length of utterance (MLU) as well as an increase in grammatical errors (Ash et al., 2006, 2009, 2010, 2013; Graham et al., 2004; Grossman, 2012; Jokel et al., 2014; Knibb et al., 2009; Rohrer et al., 2010a; Sajjadi et al., 2012; Santos-Santos et al., 2016; Tetzloff et al., 2019; Thompson et al., 1997, 2012, 2013; Wilson et al., 2010). Additionally, some individuals with nfvPPA show impaired syntax in writing (Grossman et al., 1996; Tetzloff et al., 2019). Motor speech deficits and impaired grammatical ability may present in isolation or in combination and vary in severity. Some have argued that nfvPPA without agrammatism constitutes a distinct variant, referred to as primary progressive apraxia of speech (PPAOS) (Josephs et al., 2012; Jung et al., 2013; Tetzloff et al., 2019); whereas patients with relatively pure agrammatism have been referred to as PPA-agrammatic (Mesulam et al., 2009). The range of clinical profiles encompassed by the current consensus designation of nfvPPA highlights the need for consideration of an array of deficits when designing appropriate interventions and characterizing treatment benefits for this population.

Treatment of Nonfluent Aphasia: Evidence from Stroke-Induced Aphasia Research

Research investigating speech-language interventions for nfvPPA is sparse relative to treatment for nonfluent aphasia secondary to stroke. However, the treatment literature for nonfluent stroke-induced aphasia is informative when considering treatment for nfvPPA due to some similarities in behavioral presentation. More specifically, individuals with nonfluent aphasia caused by stroke also can present with both agrammatism and motor speech deficits. As such, clinicians may utilize interventions targeting specific aspects of fluency or speech production (e.g., syntax or motor speech) or those that target linguistic and motoric domains simultaneously. Treatments addressing grammatical deficits in nonfluent aphasia include programs aimed at increasing syntactic complexity in production through explicit (e.g., Treatment of Underlying Forms: Thompson & Shapiro, 2005) or implicit training of syntactic structures (e.g., Sentence Production Program for Aphasia: Helm-Estabrooks et al., 2000) or by retraining the links between surface sentence form and underlying meaning (e.g., Mapping Therapy, Rochon et al., 2005). Other interventions address deficits in fluency at the motor and/or phonological level through motor learning via sensory feedback approaches incorporating directed articulation and phonological association practice (e.g., Sound Production Treatment: Wambaugh et al., 1998; Wambaugh & Mauszycki, 2010; Phonetic Placement Therapy: Van Riper, 1947, Austermann et al., 2008; Ballard et al., 2007). Additionally, some have targeted fluency in speech production through metrical pacing or the incorporation of melody in production of phrases or discourse (e.g., Albert et al., 1973; Wambaugh et al., 2012).

Script training is another treatment approach for nonfluent aphasia in which individuals learn a dialogue or monologue verbatim through repeated practice of scripted material. This approach has shown the potential to improve motoric and grammatical dimensions of fluency in the spoken production of people with stroke-induced nonfluent aphasia and apraxia of speech (Bilda, 2011; Cherney et al., 2008; Cherney & Halper, 2008; Goldberg et al., 2012; Grasso et al., 2019; Holland et al., 2002; Lee et al., 2009; Youmans et al., 2005, 2011). It is thought that intensive, repetitive practice of scripts lends automaticity to the act of speaking the scripted material, capitalizing on the tendency of automatic speech to be preserved in those with nonfluent aphasia and apraxia of speech (Dronkers, 1996; Lum & Ellis, 1999). Through repetition and exposure to correct grammatical structures, script training is also thought to improve syntactic production without explicit training of syntax (Cherney et al., 2008; Goldberg et al., 2012; Youmans et al., 2005). Scripts may be generic or personalized to specific areas of interest or communicative needs in order to promote functionality for the individual. Often, participants proceed through a hierarchy of training tasks including choral reading, repetition, and independent production. "Speech-entrainment," a training technique in which a participant speaks in unison with an audio-visual model of a healthy speaker, has been observed to have fluency-enhancing effects for scripted material in individuals with nonfluent aphasia (Bonilha et al., 2019; Fridriksson et al., 2012, 2015). Script training interventions implementing speech-entrainment or unison speech production tasks have resulted in improvements in production of scripted material and speech rate for people with stroke-induced nonfluent aphasia or apraxia of speech via training with a clinician (Ali et al., 2018; Bilda, 2011; Costello-Yacono & Balasubramanian, 2018; Fridriksson et al., 2012, 2015; Goldberg et al., 2012; Szabo et al., 2014; Youmans et al., 2005) or a virtual therapist model (Cherney et al., 2008, 2014, 2019; Cherney & Halper, 2008; Lee et al., 2009). Additionally, script training has been successfully applied in the treatment of nfvPPA (Henry et al., 2018), which we discuss further in the following section.

TREATMENT OF NONFLUENT/AGRAMMATIC VARIANT PRIMARY PROGRESSIVE APHASIA

In the initial stages of nfvPPA, restitutive treatment has been shown to be efficacious in improving speech and language abilities and may hold prophylactic value, protecting treated targets and skills in the face of progressive decline (Henry et al., 2018; Meyer et al., 2019; Murray, 1998). While there are numerous studies examining treatment effects for interventions targeting deficits in lvPPA and svPPA (for reviews, see Cadório et al., 2017; Croot et al., 2009; Jokel et al., 2014; Kortte & Rogalski, 2013; Rising, 2014; Tippett et al., 2015), there are only a handful of studies that have examined restitutive speech-language interventions for grammatical or motor speech deficits in nfvPPA (Hameister et al., 2017; Henry et al., 2013, 2018; Machado et al., 2014; Schneider et al., 1996). Results from these interventions indicate the potential for grammatical and motor speech treatments adapted from approaches in stroke-induced aphasia to benefit individuals with nfvPPA as well.

Some studies have focused on training verb tenses in order to address grammatical deficits in nfvPPA. In one study, a participant was trained in the production of past, present progressive, and future verb tenses through spoken production and a system of gestures. In response to questions regarding photo stimuli, either verbal or gestural responses were accepted as correct. Results indicated this yielded improved performance in sentence production with trained verbs and showed generalization to untrained verbs within trained tenses. Similarly, another study trained verb tenses in the context of sentence production through the use of a cloze procedure technique (Machado et al., 2014). The individual with nfvPPA, whose verbal output at pre-treatment was limited to nouns and short phrases, showed improvement in production of correct verb tenses in sentence completion tasks at post-treatment.

Another intervention that focused on retrieval of verbs also incorporated training of appropriate grammatical structures in verbal descriptions of simple picture stimuli (Hameister et al., 2017). The protocol was a form of constraint-induced therapy in which the two participants with nfvPPA were encouraged to speak in syntactically correct sentences or phrases and refrain from using other communicative modalities. A clinician provided feedback and modeled grammatically appropriate phrases and sentences of increasing complexity over the treatment period. Results indicated that the participants produced significantly more complete, grammatical responses to the picture stimuli at post treatment.

Other studies have addressed the motoric deficits common to individuals with nfvPPA. In a novel approach involving structured oral reading, Henry and colleagues targeted articulation of multisyllabic words in a participant with mild AOS and nfvPPA (Henry et al., 2013). While reading aloud, the clinician guided the participant through a hierarchy of speech production tasks for misarticulated multisyllabic words, beginning with segmented (syllable-by-syllable) articulation and progressing to production of the word in the context of the full sentence. Additionally, the participant was assigned oral reading homework and encouraged to self-identify errors and employ practice with the segmentation hierarchy learned in treatment sessions. Improvements were observed in the number of speech errors and successful self-corrections in reading of trained and untrained texts at post treatment. These improvements were maintained at 12-month follow-up, with continued home practice.

Another intervention targeting more severe motor speech deficits in nfvPPA was adapted from rate and rhythm control interventions for AOS in stroke-induced aphasia (Beber et al., 2018). Treatment employed pacing strategies and elongation of initial sounds in words to aid with initiation and production of speech. Qualitative results indicated increased independent implementation of the strategies, which were observed to aid in production of single words and short phrases.

Script training, as previously mentioned, holds promise in addressing the core deficits of nfvPPA. Henry and colleagues (2018) examined the results of intervention in mild-moderate nfvPPA using Video-Implemented Script Training for Aphasia (VISTA, Henry et al., 2018). VISTA is an intervention in which production of personalized scripts is trained via practice with a clinician as well as at-home training using speech entrainment practice. Videos are created for each script, with content spoken by a gender-matched healthy speaker serving as an audio-visual model. Linguistic and articulatory complexity are tailored to the individual patient and the rate at which the model produces the participant's scripts is determined on an individual basis, taking into account their current spontaneous speech rate. Targeted practice of the script in training is undertaken during twice-weekly sessions with a clinician, proceeding through a hierarchy of tasks targeting accurate speech production and memorization. Additionally, participants are required to participate in speech-entrainment home practice with their script for at least 30 minutes per day in order to promote memorization of content and automaticity of production.

Outcome measures in this initial pilot study indicated that intervention with VISTA resulted in improvements in percent correct, intelligible scripted words and reductions in grammatical errors for trained scripts as well as improved overall intelligibility at post-treatment. Importantly, improvements for trained scripts showed maintenance over a year-long follow-up period, suggesting the effects of this intervention are robust. While the target of massed script practice is learned automaticity of scripted content, the implicit goal of treatment is improved functional communicative ability. Based on the theory of automatization (Logan, 1988), it is traditionally not expected for improvement related to script training to generalize to other tasks or areas of deficit as the training involves practice of a whole task procedure rather than individual component skills. However, comprehensive analysis of changes on untrained discourse samples following VISTA and other script-training interventions has yet to be undertaken.

DISCOURSE ANALYSIS IN APHASIA

Although the use of standardized cognitive-linguistic assessment tools is common in clinical practice, these assessments are thought to have shortcomings in predicting communication abilities in functional contexts (Beeke et al., 2008; Herbert et al., 2008; Kemper & Kemper, 2006; Mayer & Murray, 2003; Ulatowska et al., 2003). Discourse analysis involves characterizing and evaluating linguistic structures from connected speech at the sentence level and beyond. This approach has been applied in aphasia treatment studies with increasing frequency (Bryant et al., 2016). Through examination of speech in tasks that more closely resemble real-world communicative contexts, a clearer picture of true abilities may emerge.

Several methods of linguistic assessment of discourse have been implemented in aphasia treatment research (e.g., Quantitative Production Analysis: Saffran et al., 1989; Shewan Spontaneous Language Analysis: Shewan, 1988; Language Assessment Remediation and Screening procedure: Crystal et al., 1976; Northwestern Narrative Language Analysis, Thompson, 2013). These systems are often used in conjunction with measures of communication efficiency and information content (e.g., calculation of correct information units, Nicholas & Brookshire, 1993) to provide a comprehensive characterization of productive language ability.

One of the more commonly applied systems, Quantitative Production Analysis (QPA, Saffran et al., 1989), has been used to derive numerous measures from discourse samples. Saffran et al. outline a reproducible method for extracting samples from productions of well-known narratives (e.g., a classic fairytale), segmentation of utterances based on a hierarchy of structural indexes, and quantification of a variety of linguistic components that can be compared across groups and individuals. Data yielded via QPA include morphological, lexical, and syntactic measures. QPA and other systems have been implemented in studies of aphasic speech, aiding in identification of measures which reliably distinguish between the speech of people with aphasia and controls (Fromm et al., 2016; Rochon et al., 2000; Saffran et al., 1989; Thompson & Shapiro, 1995), as well as describe differences between subtypes of aphasia and PPA (Ash et al., 2013; Bird & Franklin, 1996; Fraser et al., 2014; Fromm et al., 2016; Mack et al., 2015; Saffran et al., 1989; Wilson et al., 2010). However, these discourse analysis procedures frequently require elicitation of samples of a specific length, which can be prohibitive for more impaired individuals with sparse output. Additionally, this method necessitates extensive coding that is both time-consuming and demands high-level knowledge regarding syntactic structures, limiting its functional utility in standard clinical settings.

Computerized analysis tools have been developed to address the burden of transcribing and coding discourse samples by hand. Of these, Computerized Language ANalysis (CLAN, MacWhinney, 2000) and Systematic Analysis of Language Transcripts (Miller & Iglesias, 2012) are the most frequently used software applications in discourse analysis in aphasia treatment (Bryant et al., 2016). These programs allow for automated calculation of a variety of linguistic measures derived from transcribed samples coded for specific forms and structures, reducing both the time demand and potential for human error in calculation or identification of particular morphological or syntactic structures. Though analysis via CLAN has yet to prove as accurate as hand-coding methods for aphasic speech analysis at the sentence- or utterance-level without the use of additional hand-coded elements (e.g., Hsu & Thompson, 2018), the automation of coding of morphological and lexical structures holds promise for enabling discourse analysis to be implemented more efficiently in research and for improving access to and use of this diagnostic tool for clinicians.

Because script training generally targets speech production for use in a functional context, discourse analysis is an ideal candidate for capturing treatment-induced change for this intervention. As yet, a handful of script training treatment studies in nonfluent aphasia have used elements of discourse analysis (Cherney et al., 2008; Cherney & Halper, 2008; Costello-Yacono & Balasubramanian, 2018; Fridriksson et al., 2012; Goldberg et al., 2012; Grasso et al., 2019; Henry et al., 2018; Nobis-Bosch et al., 2011; Youmans et al., 2005). These studies selected a variety of measures to characterize treatment-related change in their participants, including measures of content, grammaticality, fluency, and intelligibility.

The content of connected speech samples in script training studies has generally been examined through comparison of output to the target scripted words (e.g., number of scripted words). Fridriksson et al., however, instead calculated the percentage of different words produced and found significant increases following script training with speech-entrainment (Fridriksson et al., 2012). Grammatical measures including number of script-related morphemes, nouns, verbs, and modifiers (Cherney et al., 2008; Cherney & Halper, 2008), number of grammatical errors per 100 words (Grasso et al., 2019; Henry et al., 2018), subject-verb-object structure production (Costello-Yacono & Balasubramanian, 2018), and percentage of words with grammatical morphemes (Goldberg et al., 2012) have also shown relative improvement at post-treatment. Two studies (Goldberg et al., 2012; Youmans et al., 2005) also examined disfluencies (e.g., repetitions and revisions), revealing a reduction at post-treatment that was significant in Youmans et al., 2005 but not for Goldberg et al., 2012. Intelligibility, as measured by percent of intelligible words, was found to significantly increase post-script training in two studies (Grasso et al., 2019; Henry et al., 2018). Speech rate was also found to be sensitive to script training effects, showing an increase at post-treatment (Ali et al., 2018; Cherney et al., 2008, 2014; Cherney & Halper, 2008; Costello-Yacono & Balasubramanian, 2018; Goldberg et al., 2012; Moss, 2009; Szabo et al., 2014; Youmans et al., 2005, 2011).

While these studies illuminate the potential for discourse analysis to characterize treatment-induced changes relative to script training, most focus on stroke-induced aphasia, where the potential for recovery is well documented, and very few studies have examined discourse-level changes in untrained speech samples.

MEASURING GENERALIZATION IN DISCOURSE MEASURES FOLLOWING SCRIPT TRAINING

Script training has been shown to be efficacious in the treatment of stroke-induced nonfluent aphasia and nfvPPA in that it gives individuals a means of communicating functional content in their daily life. Investigation of generalization effects in script training provides a window into whether script training could yield some benefit beyond explicitly trained scripts, addressing the underlying deficits associated with nonfluent aphasia in stroke and nfvPPA.

Some studies examining the effects of script training in stroke-induced aphasia have documented improvement at post-treatment in communicative ability as measured by functional communication assessments or self-report measures (Bilda, 2011; Cherney et al., 2011, 2015; Fridriksson et al., 2012; Goldberg et al., 2012; Manheim et al., 2009; Nobis-Bosch et al., 2011; Youmans et al., 2011). Others have investigated generalization in terms of how well the participants produced their scripts in novel contexts or with novel partners (Goldberg et al., 2012; Youmans et al., 2005). However, to our knowledge, only two studies have examined the potential for generalization of benefit to untrained discourse following script training (Costello-Yacono & Balasubramanian, 2018; Nobis-Bosch et al., 2011). In Nobis-Bosch et al., 2011, participants' responses during an interview at pre- and post-treatment were analyzed for linguistic parameters (i.e., percentage of open class words, type token ratio, percentage of syntactically complete clause-like units, percentage of clause-like units in compound sentences, and MLU) using Aachen-Sprach-Analysis (Grande et al., 2008). Results indicated that small subsets of participants improved on at least one linguistic measure (e.g., four of 18 improved with regard to MLU; two of 18 improved for percentage of open class words), but no specific patterns of change related to script training were found (Nobis-Bosch et al., 2011). Costello-Yacono and Balasubramanian, 2018, compared the effects of script training with that of Verb Network Strengthening Treatment (VNeST, Edmonds, L., Nadeau, S., Kiran, 2009) in two participants with nonfluent aphasia in a crossover study. Connected speech samples were collected once weekly during each of the interventions (each 9 weeks in duration) through elicitation of a procedural narrative, picture description, or response to a short video. These were transcribed and analyzed for speech rate (words per minute), sentence-verb-object productions, and errors (number of paraphasias, repetitions, omissions, substitutions, non-responses, incomplete utterances, morphological errors, fillers, and perseverations). Numerical increases were observed during the course of script training for one participant in speech rate and sentence-verb-object production, while errors remained variable for both participants (Costello-Yacono & Balasubramanian, 2018). Further examination of the Percentage of Data Exceeding the Median scores (PEM) for these measures indicated that script training was moderately effective for one participant for rate of speech and production of sentence-verb-object structures, highly effective for one participant for rate of speech, and ineffective for reducing errors production in both participants (Costello-Yacono & Balasubramanian, 2018).

With regard to nfvPPA, the results of Henry et al. (2018) provide an indication that script training may result in generalized improvement in grammar, as evidenced by a decrease in the number of grammatical errors in untrained script topics and improvement on a standardized test of grammatical production (the Northwestern Anagram Test, Weintraub et al., 2009) for some participants; however, significant changes were not observed at the group level for these metrics. Further investigation is needed in order to assess the potential for script training to benefit grammatical ability and fluency in untrained connected speech in both stroke-induced nonfluent aphasia and nfvPPA.

To our knowledge, no script training studies have characterized treatment response via detailed linguistic analysis examining multiple language domains in both trained and untrained language samples in nfvPPA. Given preliminary evidence that measures of speech production and fluency may be sensitive to treatment-induced change via script training in nfvPPA, a more thorough investigation of candidate outcome measures derived from discourse analysis and applied to both trained and untrained language samples is warranted.

CURRENT STUDY

In the pilot VISTA study, accuracy, intelligibility and grammaticality of production were analyzed, with promising outcomes for trained scripts and minimal generalization to untrained content. In the current study, we evaluated the utility of additional, largely-automated discourse analysis procedures for characterizing VISTA treatment response. Specifically, we examined treatment-induced changes on metrics capturing speech fluency, grammar, and informativeness at the discourse level in a larger nfvPPA patient sample (n=10 from the original study and n=10 new patients). Variables of interest were extracted from participants' trained and untrained connected speech samples using CLAN software and performance at post-treatment was compared to pre-treatment.

Based on findings from the previous study (Henry et al., 2018), we predicted that trained script samples would show significant improvement on discourse measures from pre- to post-treatment. Specifically, we predicted a significant increase in words per minute, mean length of utterance in morphemes, grammatical complexity, and propositional density and a significant decrease in fluency disruptions per hundred words and proportion of open to closed class words. We also predicted that improvements would differ significantly between trained and untrained topics from pre- to post-treatment, with trained topics demonstrating greater improvement. Finally, we predicted that generalized improvement on untrained samples would not be significant at the group level, but that individual participants may show evidence of generalized improvement in discourse.

Methods

Participants

A total of 20 individuals (12 females) meeting current consensus criteria (Gorno-Tempini et al., 2011) for nonfluent/agrammatic variant PPA participated in this study. Participants were recruited via the primary research sites (the University of Texas at Austin or the University of California, San Francisco) where written consent to participate was obtained from all individuals. All procedures were approved by the institutional review boards at both institutions. The cohort had a mean age of 68 (range: 57-78), and an average of 17 years of education (range: 12-22). All participants were White/Caucasian native speakers of English and were functionally monolingual. Assessment data and data regarding treatment outcomes were previously reported for a subset of participants included in this study (n = 10; Henry et al., 2018).

To be considered eligible for participation in the current study, individuals were required to meet diagnostic criteria for nonfluent/agrammatic variant PPA including the core features of agrammatism in production and/or effortful, halting speech, consistent with features of apraxia of speech (Gorno-Tempini et al., 2011). Diagnosis of PPA variant was determined by neurologists employing current consensus criteria and verified by comprehensive speech-language and cognitive assessment conducted prior to initiation of treatment (see Table 1). All participants demonstrated motor speech impairment with features of apraxia, and all but six presented with features of dysarthria as noted during a motor speech evaluation (Wertz et al., 1984). Fifteen participants showed impaired expressive grammar in connected speech and/or on the Northwestern Anagram Test (NAT, Weintraub et al., 2009), and five participants showed minimally impaired expressive grammar on standardized testing (score of 90% or greater on the NAT) and minimal to no grammatical impairment in connected speech. In addition to meeting diagnostic criteria for nfvPPA, participants were required to attain a score of 15 or greater on the Mini-Mental State Exam (MMSE, Folstein et al., 1975), and demonstrate intact repetition of at least five syllables on the repetition subtest of the Western Aphasia Battery-Revised (WAB-R, Kertesz, 2006), at pre-treatment. The average pre-treatment MMSE score was 27 (range: 23-30; SD 2.4), indicating relatively spared cognition. In general, individuals were mildly aphasic at pre-treatment as indicated by the average Aphasia Quotient on the WAB-R (86.4; range: 65-97.2; SD 9.0). For additional details regarding demographics and cognitive and linguistic assessment scores for participants, see Table 1.

Participation in the study also required either access to a stable internet connection in order to participate via teletherapy or the ability to attend in-person sessions at either of the research sites. Lastly, individuals with significant uncorrected hearing or visual impairment were not eligible for participation in the study.

PROCEDURE

All participants were treated using Video-Implemented Script Training for Aphasia (VISTA, Henry et al., 2018) as part of an ongoing PPA intervention study. Treatment was administered using a single-subject multiple baseline design. Participants were seen by a clinician in person (n = 6) or via HIPAA-compliant videoconferencing software (n = 14; Zoom© or Fuze©) if their distance from the treatment sites precluded in-person visits.

Script Development and Stimuli Creation

Prior to the initiation of treatment, participants provided the clinician with six conversational topics relevant to their individual communication interests and needs (e.g., “family” or “button collection”). The clinician then worked in conjunction with the participant to develop scripts for each of the six topics. Script length varied from four to seven sentences. The level of grammatical and motor speech impairment of the individual participants was taken into account when developing the scripts, yielding scripts which would challenge their abilities while still serving as feasible targets. The scripts were balanced for number of words, number of sentences, number of complex words (words with three or more syllables), mean words per sentence, mean syllables per word, as well as readability using Flesch Kincaid reading scores (Flesch, 1948). Four scripts were randomly selected for training while the remaining two served as untrained controls. Prior to finalization, the scripts were shown to the participants once in order to ensure accuracy of content and to confirm that their composition was congruent with the participant's individual communicative style.

Video stimuli were created for each of the six scripts by recording the mouth of a healthy speaker, who spoke the script at a steady rate with exaggerated articulatory gestures. The speaking rate at which the scripts were recorded was determined by consideration of the participant's speech rate during oral reading and picture description tasks. Ten participants were treated with the VISTA protocol with the addition of rate manipulation (VISTA-RM), a modification to the original treatment procedures that was designed to adaptively increase practice difficulty, as appropriate, based on patient performance. For these participants, videos for each script were created at two additional speaking rates by increasing the rate of the video using video and audio editing software (Adobe After Effects and Adobe Audition). This yielded a set of script videos at the participant's "starting rate," which was determined by the same procedure as the original VISTA protocol, as well as sets of script videos 10% and 20% faster than the participant's starting rate. The instances in which these increased practice rates were implemented in treatment are described below.

Study Design and Treatment Protocol

After finalization and recording of the scripts, the participants were asked to speak about each of the scripted topics during two pre-treatment probe sessions in order to establish baseline performance. The clinicians asked the participants to "Tell me about [the script topic]" and instructed them to try to remember the script they had developed together as best they could, having only seen the finalized scripts once.

A practice video for each script was provided to the participants for the interval in which it was actively in treatment. The participants were instructed to practice with their script video for at least 30 minutes per day. Home practice consisted of watching the video with headphones on a computer or iPad provided by the researchers (for the duration of the study) and attempting to speak in unison with the healthy model. Homework also included targeted articulatory practice of a few scripted words or short phrases which the participant had struggled with during the previous treatment session, as appropriate.

Treatment consisted of two 45-minute to 1-hour sessions per week. Sessions began with the collection of probes for the script in training and two of the other five scripts, with all scripts probed at least once each week. For the ten participants who were administered the original VISTA protocol, each script was trained for either 2 or 3 sessions, depending on performance. During probes, the number of correct, intelligible scripted words produced for each topic was calculated. A word was counted as correct if it was present relative to the words in the script and intelligible within the context of the scripted topic. If the participant met the criterion of 90% correct, intelligible scripted words during the probe at the beginning of the second session, a new script would enter treatment the following session. The participant then began practicing the new script in homework and in sessions with the clinician; the previously trained script topic would continue to be probed each week but was not practiced further. For the ten participants engaged in the VISTA-RM protocol, all scripts were trained for 3 sessions. The clinician would consider the participant's performance in both the spontaneous probe at the start of the session and a unison speech probe with the current script video in order to determine whether the 90% criterion was met. If criterion was reached in both conditions, the participant was provided with a new home practice video with speaking rate increased by 10% in order to enhance practice difficulty.

During sessions with the clinician, a hierarchy of tasks ranging from more structured to more functional activities (see Table 2) was utilized to promote memorization and conversational usage of scripts. For instance, the participants were asked to select scripted sentences from amongst foils and engage in targeted articulation practice with words and phrases in order to maximize intelligibility. Functional tasks included responding to questions from the clinician and using scripted sentences in the context of a conversation with a novel listener.

Immediately following the end of treatment, participants again underwent speech-language and cognitive testing, and completed two sets of script probes during which all trained and untrained scripts were elicited.

Transcription Procedure

All pre- and post-treatment script probe time points (4 samples in total, 2 pre-treatment and 2 post-treatment) were recorded either via digital camera if in person or via teleconferencing software. Using these recordings, script probes were transcribed by speech-language pathology undergraduate or graduate students in the Aphasia Research and Treatment Lab who were trained in transcription and coding procedures.[1]

For each participant, one time point was randomly selected to be transcribed a second time by one of the graduate students (K.B., K.S., or W.K-R.) in order to evaluate reliability. Transcriptions were then coded for analysis in CHAT (MacWhinney, 2000), the transcription program accompanying CLAN, by a single graduate research assistant (K.B.) to ensure consistency. Discrepancies in utterance boundaries between the two transcribers were resolved by consensus using criteria for utterance segmentation established for QPA by Saffran et al., 1989. After resolving utterance-level discrepancies, the transcriptions were compared at the word level using CLAN. The average percentage of matching words between the two transcribers for all dual-transcribed timepoints was 94%. After reliability was calculated, discrepancies in words were resolved via consensus.

Measuring Overall Script Production Accuracy

Percent of correct, intelligible scripted words at the pre- and post-treatment timepoints was collected by the clinicians for each script (trained and untrained). This measure served as the primary outcome measure in Henry et al., 2018, which reported on treatment effects for 10 of the nfvPPA participants included in the current study. While this measure was not the focus of the current study, documenting change on this measure provides information regarding the efficacy of the intervention approach in this larger

[1] Four script probes were transcribed for all but two participants, where the audio quality of the recordings precluded a subset of probes from being accurately transcribed.

sample. As such, it provides context for interpreting change in performance on the novel discourse metrics evaluated in the current study.

DISCOURSE MEASURES

Coded transcriptions were analyzed using CLAN for several measures, as detailed below. These measures were chosen in order to evaluate the content, syntactic complexity, and rate of the participants' speech based on prior research documenting connected speech impairments in nfvPPA (Ash et al., 2006, 2009, 2010, 2013; Croot et al., 2012; Graham et al., 2004; Grossman, 2012; Jokel et al., 2014; Knibb et al., 2009; Rohrer et al., 2010a; Sajjadi et al., 2012; Santos-Santos et al., 2016; Tetzloff et al., 2019; Thompson et al., 1997, 2012, 2013; Wilson et al., 2010) as well as prior research examining changes in speech production and fluency following script training in stroke-induced aphasia (Ali et al., 2018; Bilda, 2011; Cherney et al., 2008, 2014, 2019; Cherney & Halper, 2008; Costello-Yacono & Balasubramanian, 2018; Goldberg et al., 2012; Grasso et al., 2019; Holland et al., 2002; Lee et al., 2009; Moss, 2009; Nobis-Bosch et al., 2011; Szabo et al., 2014; Youmans et al., 2005, 2011). Given the time-intensive nature of transcription and analysis of connected speech, another aim in selecting the following measures was to limit the amount of hand-coding necessary to analyze a large number of discourse samples.

In contrast with traditional methods of discourse analysis, (e.g., Quantitative Production Analysis), the entirety of a participant's response to a script probe was transcribed and entered into analysis, with the exclusion of only clearly off-topic comments. Whereas, for instance, Quantitative Production Analysis excludes habitual starters (e.g., "and then"), coordinating conjunctions joining utterances, and direct discourse markers (e.g., “he said, ‘X’”) from the sample, we did not. Our more liberal approach to inclusion was chosen due to the open-ended nature of the probes and a desire to avoid subjective decisions regarding the exclusion of utterances. For our purposes, off-topic comments subject to exclusion were defined as asides addressing an intervening event during the participant’s response (e.g., “my phone is ringing” or “they’re mowing outside”) or clear comments on the task (e.g., “this one is hard”).

The CLAN program parses morphological and grammatical information using natural language processing algorithms. After entering transcriptions with utterances separated onto individual lines, the programs MOR, PREPOST, POST, POSTMORTEM, and MEGRASP may be run using a single command. This series of programs parses morphemes and disambiguates the grammatical relationships amongst them, producing tiers in the transcript which mark each part of speech and indicate their syntactic relationships. This process was performed on all the scripts prior to running further analysis. In the following sections, we provide details regarding each of the measures of interest in the current study.

Words per Minute (WPM)

Words per minute was chosen as the speech of individuals with nfvPPA has been documented to be markedly reduced in rate (Ash et al., 2010, 2013; Croot et al., 2012; Santos-Santos et al., 2016; Thompson et al., 1997, 2012; Wilson et al., 2010). Additionally, evidence from studies in stroke-induced nonfluent aphasia indicates that script training has a beneficial effect on speech rate (Ali et al., 2018; Cherney et al., 2008, 2014; Cherney & Halper, 2008; Costello-Yacono & Balasubramanian, 2018; Goldberg et al., 2012; Moss, 2009; Szabo et al., 2014; Youmans et al., 2005, 2011). Duration information was recorded by timing the length, in seconds, of the participant's response after the clinician's probe, "Tell me about [script topic]." The start time was indicated as beginning precisely with the participant's first utterance, with the exception of a restatement of the topic in isolation or a habitual response to the probe (e.g., "Okay"), where timing commenced with the first word immediately following. The end time corresponded with the end of the final word spoken before the participant indicated they were finished. In some instances, at the end of the probe, the clinician would ask, "Is there anything else you would like to say?" If the participant then added more to their response, this was included in the transcription, with the duration of the clinician's remark subtracted from the total duration. The count of total words was derived by summing counts of intelligible and unintelligible words performed in CLAN with the FREQ program. In order to calculate words per minute, the count of total words in a probe, including unintelligible words and words contained within a repetition or revision, was divided by the total duration, which was then multiplied by 60.

Fluency Disruptions per Hundred Words

Evidence from previous studies reveals that connected speech in nfvPPA is characterized by repetitions, revisions, and phonological paraphasias (Ash et al., 2010, 2013, 2019; Croot et al., 2012; Santos-Santos et al., 2016; Thompson et al., 1997, 2012; Wilson et al., 2010). Additionally, studies of script training in stroke-induced nonfluent aphasia have shown reductions in repetitions and revisions in connected speech following treatment (Goldberg et al., 2012; Youmans et al., 2005). Given the number of different elements which may disrupt fluency in individuals with nfvPPA, we endeavored to capture all of these in a measure balanced for the length of the sample. Our measure was a sum of the following four components, which represent disruptions to speech fluency: number of fillers, number of phonological fragments, number of repetitions, and number of revisions. This sum was then divided by the total number of words produced, as counted by summing counts of intelligible and unintelligible words in CLAN with the FREQ program, then multiplied by one hundred. All of the following were coded with specific markers in the transcriptions and then automatically summed through CLAN.

Number of Fillers

Fillers were indicated as defined by the CLAN manual: "uh," "um," "er," "eh," "you know," with the addition of "like." "You know" and "like" were only counted as fillers where they were habitual or they were clearly not acting as a meaningful structure.

Number of Phonological Fragments

A phonological fragment was defined as one or several phonemes produced together, which did not form a complete recognizable word in context. Successive approximations of a target word, as seen in a conduit d'approche, that were not intelligible as the target word in context were also considered fragments for coding purposes. If a clear conduit d'approche was initiated and did not reach an intelligible target at any point, the final or most complete string of phonemes was coded as an unintelligible word and the rest of the attempts were coded as phonological fragments. Strings of phonemes that were not clearly successive attempts at a single target or abrupt phonological fragments were coded as unintelligible words, and therefore not included in the sum of fluency disruptions per hundred words.

Number of Repetitions

Repetitions were words or phrases that were repeated verbatim.

Number of Revisions

A revision was counted when the participant produced a word or phrase, then altered its lexical content, syntax, or pronunciation while maintaining the same conceptual information.

Mean Length of Utterance (MLUm)

Mean length of utterance was measured in morphemes. This measure was included because previous studies have documented reductions in mean length of utterance in the speech of individuals with nfvPPA (Ash et al., 2006, 2009, 2010, 2013; Graham et al., 2004; Grossman, 2012; Jokel et al., 2014; Knibb et al., 2009; Rohrer et al., 2010b; Sajjadi et al., 2012; Santos-Santos et al., 2016; Tetzloff et al., 2019; Thompson et al., 1997, 2012, 2013; Wilson et al., 2010). Additionally, research in stroke-induced aphasia has shown changes in MLU (Nobis-Bosch et al., 2011) and numbers of script-related morphemes produced (Cherney et al., 2008; Cherney & Halper, 2008) following script training. Revisions and repetitions were excluded from this measure. Utterances containing unintelligible words were included; however, the unintelligible words themselves were not counted. This measure was automatically calculated with the MLU program in CLAN.

Grammatical Complexity Index

In individuals with nfvPPA who present with grammatical deficits, connected speech is characterized by reductions in syntactic complexity (Ash et al., 2010, 2013; Graham et al., 2004, 2016; Knibb et al., 2009; Rogalski et al., 2011; Sajjadi et al., 2012; Tetzloff et al., 2018, 2019; Wilson et al., 2010). Research in nonfluent stroke-induced aphasia provides evidence that script training improves production of grammatical morphemes for trained material (Cherney et al., 2008; Cherney & Halper, 2008; Goldberg et al., 2012) and production of more complete sentences (i.e., sentence-verb-object structures, Costello-Yacono & Balasubramanian, 2018). However, most analyses of complexity of syntax in connected speech require extensive hand-coding. We endeavored to examine the feasibility of using an automatically calculated measure of grammaticality to characterize treatment-related change following script training. CLAN allows for relatively automatic calculation of such a measure in the form of the grammatical complexity index. The grammatical complexity index is determined by counting the number of grammatical relations that are markers of syntactic embeddings and dividing this sum by the total number of grammatical relations in a sample. According to the work of Kimberly Mueller, who formulated the procedure in CLAN, the accuracy of this process using CLAN is about 95%, which is comparable to that of human coding (MacWhinney, 2000). The grammatical complexity index is calculated through division of values provided by CLAN through the FREQ program.

Propositional Idea Density

In selecting propositional idea density, we hoped to investigate its utility as a measure of treatment related change in informativeness in connected speech that did not require the extensive hand-coding associated with Correct Information Unit (CIU) analysis. Propositional idea density (sometimes referred to as "idea density") describes the density of information conveyed, or propositions, in speech by summing the number of verbs, adjectives, adverbs, prepositions, and conjunctions and dividing by the total number of words. Research has shown that individuals with stroke-induced nonfluent aphasia present with reduced propositional density as compared to control subjects (Bryant et al., 2013; Ferguson et al., 2013; Fromm et al., 2016; Ulatowska et al., 1981, 1983) and other subtypes of aphasia (Fromm et al., 2016). To our knowledge, propositional density of connected speech in nfvPPA has only been examined in one study, which found it to be effective in distinguishing nfvPPA from lvPPA and svPPA, with nfvPPA associated with decreased propositional density (Vander Woude, 2017). Additionally, propositional density has been shown to be a sensitive predictor of Alzheimer's disease and decline in language associated with aging (Butler & Snowdon, 1996; Kemper et al., 2001; Mortimer & Borenstein, 2012; Riley et al., 2005; Snowdon et al., 1996). No hand coding is required for calculation of propositional idea density in CLAN. CLAN's propositional idea density measure was adapted from Computerized Propositional Idea Density Rater, third major version (CPIDR3, Brown et al., 2008).

Proportion of Open to Closed Class Words

This measure was primarily an exploratory measure. Results from studies examining nfvPPA are mixed, with some indicating a reduction in closed class words for at least a subset of participants (Ash et al., 2010; Thompson et al., 1997; Wilson et al., 2010) and others showing reductions in open class words (Ash et al., 2010). In studies of script training in stroke-induced nonfluent aphasia, only Nobis-Bosch and colleagues have examined the percentage of open class words produced in connected speech, finding that two of 18 participants improved and one declined at post-treatment (Nobis-Bosch et al., 2011). In keeping with our goal of exploring measures which did not require hand-coding, the proportion of open to closed class words can be automatically calculated via CLAN. It was included to examine its utility in characterizing treatment-related change on a morphological level, with the potential to inform changes in grammatical structures.

Open class words include all nouns, adverbs, adjectives, and participles, as well as all verbs with the exclusion of auxiliaries, copulas, and modals. Closed class words include all other parts of speech not designated as open class, and are commonly referred to as function words. The total number of open class words in a sample was divided by the total number of closed class words in a sample in order to produce this proportion. This calculation was performed in CLAN, which automatically categorizes words into open and closed class.

Statistical Analysis

For each of the aforementioned outcome measures, each participant's average for each script from pre-treatment (two observations per script at each timepoint, except for two individuals who only had one usable pre-treatment video), and post-treatment (two observations, except for one individual who only had one usable post-treatment video) were calculated and subsequently used in our analyses. We used a series of mixed-effects linear regression models with a fixed effect of timepoint (pre or post-treatment) and a random intercept for participant using the nlme package (Pinheiro et al., 2020) in RStudio version 3.5.2 (R Core Team, 2018). Analyses for paired data were conducted for the following dependent variables: MLUm, WPM, propositional density, fluency disruptions per hundred words, grammatical complexity and the ratio of open to closed class words, for trained and untrained scripts from pre- to post-treatment. In order to contextualize these outcomes, the same analysis was performed on the primary treatment outcome measure of correct, intelligible scripted words for trained and untrained scripts from pre- to post-treatment; for this variable, participants' average accuracy was calculated across trained scripts or untrained scripts at each timepoint.

To compare performance between trained and untrained content from pre- to post-treatment, we conducted additional mixed-effects linear regression models with an interaction term of time (pre and post-treatment) and condition (trained and untrained) and a random effect of participant with the same dependent variables, allowing us to infer the specificity of observed training effects.

For each analysis, residuals were assessed for outliers (who were subsequently removed from the analysis). General conformity to the assumption of normality was assessed via Shapiro-Wilk tests and quantile-quantile plots, and other assumptions were assessed via typical diagnostic methods. Because we predicted significant improvement on trained scripts and significantly greater improvement for trained scripts relative to untrained scripts, these contrasts were assessed via one-tailed tests. Two-tailed tests were used to assess performance on untrained scripts as these effects were less predictable.

Group-level statistics comprise our primary analyses. Nevertheless, because patient populations are often heterogenous in presentation and in their response to treatment, we report individual change scores (post-treatment – pre-treatment mean) for each outcome measure in order to quantify individual-level change. In order to evaluate the magnitude of changes on untrained samples at the individual level, we also identified the number of participants who showed improvement that was equal to or greater than the average improvement observed for trained script samples (for measures that showed significant improvement at the group level).

Results

Performance on Script Production Accuracy Measure Following VISTA

In order to contextualize the specific measures of interest in this study, we provide the results of a linear mixed-effects model evaluating performance on the treatment outcome measure of percent correct, intelligible scripted words. This analysis revealed a significant improvement from pre- to post-treatment ($\beta = 51.08$, $F(1,19) = 189.43$, $p < .0001$, 95% CI [43.51- 58.65]; pre-tx $M = 37.55$ % correct, post-tx $M = 88.64$ % correct, see Table 3 and Figure 1). This is consistent with data reported in Henry et al., 2018, wherein all 10 participants demonstrated a significant improvement in percent correct, intelligible scripted words for trained scripts following treatment. In contrast to the previous study, significant improvement on this measure was also observed for untrained topics ($\beta = 6.46$, $F(1,19) = 6.96$, $p = .02$, pre-tx $M = 37.89$ % correct, 95 % CI [1.47- 11.45], post-tx $M = 44.35$ % correct, see Table 3 and Figure 1).

Words per Minute[2]

Participants' performance on trained topics showed significant increases in words produced per minute from pre- to post-treatment ($\beta = 14.44$, $F(1,138) = 36.96$, $p < .0001$, 95% CI [9.77-19.11], pre-tx $M = 56.43$ WPM, post-tx $M = 71.18$ WPM; see Table 3 and Figure 2). The average change score for WPM on trained topics was 14.54 (range = -58.49-43.30), with 17 out of 20 individuals demonstrating numerical improvement (Table 4).

[2] For this measure and all others following, data and analyses are reported after removal of outliers. Change scores were calculated from average performance at the participant level.

Untrained topics, however, did not show a significant change in words produced per minute from pre- to post-treatment (β= 2.90, F (1,58) = 1.35, p = .25, 95% CI [-2.03-7.83], pre-tx M = 55.48, post-tx M = 59.25; see Table 3 and Figure 3). The average change score for WPM on untrained topics was 3.09 (range = -21.54- 21.70), with 15 individuals demonstrating numerical improvement but only three of those individuals showing change on untrained topics that was equal to or greater than the mean change observed for trained topics (see Table 4).

FLUENCY DISRUPTIONS PER HUNDRED WORDS

Significant differences were found from pre- to post-treatment in the number of fluency disruptions per hundred words that occurred when participants produced trained topics (β = -13.50, F (1,139) = 82.09, p < .0001, 95% CI [-16.43- -10.57], pre-tx M = 22.19 disruptions, post-tx M = 8.69 disruptions; see Table 3 and Figure 2). The average change score for fluency disruptions per hundred words on trained topics was -13.50 (range = -40.30-24.46), with 17 individuals demonstrating numerical improvement (see Table 4).

Performance on untrained topics also showed significant improvement from pre- to post-treatment in the number of fluency disruptions per hundred words produced (β = -3.50, F (1,59) = 4.25, p = .04, 95% CI [-6.84- -0.15], pre-tx M = 22.17 disruptions, post-tx M = 11.05 disruptions; see Table 3 and Figure 3). The average change score for fluency disruptions per hundred words on untrained topics was -3.50 (range = -22.26-21.45), with 13 individuals demonstrating numerical improvement, but only two of those individuals showing change that was equal to or greater than the mean reduction observed for trained topics (see Table 4).

Mean Length of Utterance

Participants demonstrated significant improvement for trained topics from pre- to post-treatment in MLUm (β= 1.57, F (1,139) = 25.01, $p < .0001$, 95% CI [0.95-2.18], pre-tx M = 9.82, post-tx M =11.39; see Table 3 and Figure 2). The average change score for MLUm on trained topics was 1.57 (range = -2.07-6.15), with 14 individuals demonstrating numerical improvement (see Table 4).

A nonsignificant improvement was observed for untrained scripts from pre- to post-treatment in MLUm (β= .53, F (1,59) = 1.11, p = .21, 95% CI [-0.30-1.35], pre-tx M = 10.00, post-tx M = 10.52; see Table 3 and Figure 3). The average change score for MLUm on untrained topics was .53 (range = -5.24- 3.73), with 15 individuals showing numerical improvement and six of those individuals showing change on untrained topics that was equal to or greater than the mean change observed for trained topics (see Table 4).

Grammatical Complexity Index

Participants demonstrated significant improvement from pre- to post-treatment in grammatical complexity for trained topics (β= 0.017, F (1,139) = 7.56, p = 0.004, 95% CI [0.004-.02], pre-tx M = .07, post-tx M = .09; see Table 3 and Figure 2). The average change score for grammatical complexity on trained topics was .017 (range = -0.049-0.051), with 16 individuals demonstrating improvement (see Table 4).

Participants did not, however, demonstrate a significant difference in grammatical complexity for untrained topics from pre to post-treatment (β= .006, F (1,59) = 1.25, p = .39, 95% CI [-0.007-0.02], pre-tx M = .07, post-tx M = .07; see Table 3 and Figure 3). The average change score for grammatical complexity on untrained topics was .006 (range = -

0.02-0.04), with 11 individuals showing a numerical increase and five of those individuals demonstrating improvement that was equal to or greater than the average improvement observed on trained topics (see Table 4).

Propositional Idea Density

Participants did not show significant improvement in propositional idea density from pre- to post-treatment on trained ($\beta = 0.002$, $F\,(1,139) = 0.05$, $p = .41$, 95% CI [-0.02-0.02], pre-tx $M = .46$, post-tx $M = .47$) or untrained topics ($\beta = -0.01$, $F\,(1,59) = 0.80$, $p = .37$, 95% CI [-0.03-0.01], pre-tx $M = .46$, post-tx $M = .45$).

Proportion of Open to Closed Class Words

A significant difference was observed from pre- to post-treatment in the proportion of open to closed class words for trained topics ($\beta = 0.09$, $F\,(1,137) = 4.25$, $p = .02$, 95% CI [0.004-0.17], pre-tx $M = .98$, post-tx $M = 1.07$; see Table 3 and Figure 2). The average change score for proportion of open to closed class words for trained topics was .05 (range = -.83 – 0.43), with 17 individuals showing a numerical increase (see Table 4).

Untrained topics did not show a significant change from pre- to post-treatment in the proportion of open to closed class words ($\beta = -0.05$, $F\,(1,59) = 0.47$, $p = .50$, 95% CI [-0.19-0.09], pre-tx $M = 1.02$, post-tx $M = .97$; see Table 3 and Figure 3). The average change score for untrained topics for proportion of open to closed class words was -.05 (range = -1.19 – 0.43), with 10 individuals showing a numerical increase and eight of those individuals showing change that was equal to or greater than the mean change observed for trained topics (see Table 4).

Specificity of Treatment Effects

In order to discern the specificity of treatment effects, we conducted additional analyses for all outcome measures including the interaction term of training condition (trained and untrained) and time (pre and post). For percent correct, intelligible scripted words, we observed a significant interaction between training condition and time (β= -44.62, F (1,57) = 88.61, $p < .0001$, 95% CI [-53.87--35.37]; see Figure 4), such that trained topics improved to a greater degree from pre to post-treatment relative to untrained topics, providing evidence of the specificity of the treatment effect (and experimental control) despite some generalization to untrained topics.

We now turn to the discourse measures of interest. We observed a significant effect for the interaction of training condition and time on MLUm (β= -1.04, F (1,217) = 4.20, p = .03, 95% CI [-2.11- -0.02]; see Figure 4), words per minute (β= -11.35, F (1,215) = 8.38, p = .003, 95% CI [-19.26- -3.44]; see Figure 4), fluency disruptions per hundred words (β = 10.01, F (1,217) = 15.28, $p < .0001$, 95% CI [5.00-15.01]), and the ratio of open to closed class words (β= -0.13, F(1, 215) = 3.08, p = .04, 95% CI [-0.29-0.16]) with a greater degree of improvement observed for trained topics following treatment. The interaction term was not significant for the outcome measure of grammatical complexity (β= -0.01, F(1, 217) = 1.27, p = .13, 95% CI [-0.03-0.008]), indicating that participants showed some improvement for both trained and untrained topics from pre- to post-treatment.

Summary of Results

In summary, a significant change was observed from pre- to post-treatment on the original VISTA study outcome measure of production of percent correct, intelligible scripted words. Additionally, the new discourse measures utilized for this study revealed significant changes in WPM, fluency disruptions per hundred words, MLUm, grammatical complexity, and proportion of open to closed class words. For untrained topics, a significant change was also observed in production of percent correct, intelligible scripted words and on the discourse measure of fluency disruptions per hundred words. Treatment-induced changes in propositional density were not observed for trained or untrained topics from pre- to post-treatment.

Discussion

To our knowledge, this is the first study to evaluate changes in discourse metrics following script training in individuals with nfvPPA. While studies examining the effects of script training in both nfvPPA and stroke-induced nonfluent aphasia have shown treatment-related improvements, the nature of improvements in connected speech has not been comprehensively investigated in individuals with nfvPPA.

Given the heterogeneity of behavioral presentations in individuals with nfvPPA, treatment approaches addressing the full range of possible phenotypes involving grammatical and motor speech deficits hold greater promise for broad application. As indicated by improvement on the treatment measure of percent correct, intelligible scripted words, the participants in this study experienced direct benefit from script training for the trained topics. Complementing and extending previous findings (Henry et al., 2018), improvements were also observed for trained topics on additional measures examining speech rate (i.e., WPM), speech fluency (i.e., fluency disruptions per hundred words), and grammar (i.e., MLUm and grammatical complexity). Findings indicate that script training, via repeated rehearsal of naturalistic, connected speech, successfully targets the range of impairments observed in the heterogeneous nfvPPA syndrome, including deficits in grammar and/or motor speech (i.e., apraxia of speech).

Performance on Discourse Measures: Trained Topics

Consistent with our hypothesis, participants showed significant improvements from pre- to post-treatment on trained script topics in WPM, which aligns with findings from stroke-induced nonfluent aphasia script training studies (Ali et al., 2018; Cherney et al., 2008, 2014; Cherney & Halper, 2008; Goldberg et al., 2012; Moss, 2009; Szabo et al., 2014; Youmans et al., 2005, 2011). In addition, a significant decrease in the number of fluency disruptions per hundred words at post-treatment relative to pre-treatment for trained topics was found, which has been observed in studies evaluating script training in stroke-induced aphasia (Goldberg et al., 2012). These results may be best attributed to the nature of the script training intervention, the outcome of which yields memorized, relatively automatic segments of speech. This may preclude the need for revision, repetition, or filler words resulting from difficulty with assembly of syntactic structures or with motor execution. The increase in WPM from pre- to post-treatment for trained topics indicates that repetitive practice of trained scripts may have reduced the motoric and/or linguistic demands which typically affect speech rate for participants with nfvPPA.

Additionally, in accordance with our hypothesis, MLUm and grammatical complexity were also observed to significantly improve for trained topics from pre- to post-treatment. While studies have not examined grammatical complexity specifically in the context of script training in nfvPPA or stroke-induced aphasia, these results align well with studies showing an increase in script-related morphemes (Cherney et al., 2008; Cherney & Halper, 2008). Benchmarks for meaningful change with regard to grammatical complexity, as measured by CLAN, have yet to be established. While the significant improvement observed in this study indicates promise that script training may mitigate grammatical deficits, further research is needed to determine whether this change is clinically significant. With regard to MLU, Nobis-Bosch and colleagues found increases for a small subset of individuals through analysis of interviews with participants pre- and post-treatment (Nobis-Bosch et al., 2011). However, their results failed to reveal an effect specific to script training. As yet, no other studies have directly examined MLU as it relates to treatment effects in script training for individuals with nfvPPA or stroke-induced aphasia. Our results indicate that script training allowed participants with nfvPPA to produce more complex syntactic structures and longer utterances as a result of intervention.

Contrary to our hypothesis, a significant difference was not observed in trained topics from pre- to post-treatment for propositional density. Propositional density was selected as a candidate measure for characterizing improvement in the speech of individuals with nfvPPA as we thought VISTA could positively affect informativeness. However, propositional density may have limitations as a measure of treatment-related change in connected speech in nfvPPA. Because propositions are counted based on the number of verbs, adverbs, adjectives, prepositions, and conjunctions in a sample, the number of propositions in a given sample may vary greatly depending on lexical selection. Complex verb phrases may be counted as single propositions despite being more syntactically complex than single verbs, meaning the count could be overlooking changes in complexity of verb structures. Notably, nouns are not counted in the measure of propositional density, which may disadvantage participants whose speech is primarily telegraphic (typically composed of nouns primarily). More specifically, this may not capture granular change in individuals who may have produced more unique nouns at post-treatment, effectively increasing the informativeness of their speech without the use of other word classes. Sample length also has an effect on the measurement of propositional density (Ferguson et al., 2013; Spencer et al., 2015), suggesting that, in the case of individuals with very limited output, it may not be appropriate to compare with others who produce significantly more content.

In contrast to our hypothesis, a significant increase was observed in the proportion of open to closed class words. Whereas we had expected an increase in the production of closed class words following script treatment, as they contribute to the formation of grammatical structures, the opposite pattern emerged. This may reflect a commensurate increase in the production of nouns and verbs above and beyond the increase in function words. Given that individuals with nfvPPA also exhibit word-finding difficulties, it is possible that script training's benefits are more clearly observed in increasing access to script-related words in general. Additionally, norms for individuals with nfvPPA and healthy speakers have yet to be established for this metric. Results from Wilson et al., 2010 indicated that only five of 14 participants with nfvPPA had reduced proportions of closed class words as compared with controls (Wilson et al., 2010). Other studies examining the proportion of open to closed class words in discourse tasks in nfvPPA relative to healthy controls have observed means for controls between 0.93 and 1.21 (Fraser et al., 2014; Thompson et al., 1997, 2012, 2013). For individuals with nfvPPA in these studies, the means for proportion of open to closed class words were 1.03 (Thompson et al., 2012), 1.06 (Thompson et al., 2013), and 1.09 (Fraser et al., 2014), which align well with our observed means (0.98 trained pre-tx, 1.02 untrained pre-tx; 1.07 trained post-tx, 0.97 untrained post-tx). However, in a longitudinal study of discourse in nfvPPA, the proportion of open to closed class words varied considerably between participants at various observations, ranging from 0.47 to 5.17 (Thompson et al., 1997). Notably, three out of the four participants in this study demonstrated an increase in this ratio over several years, while one individual demonstrated relative stability in the measure over seven years. This indicates that, while the predominant pattern in nfvPPA is an increase in proportion of open to closed class words with disease progression, there may be other factors influencing this metric. It is likely that examination of this measure in subsets of individuals with nfvPPA who present with a greater degree of agrammatism may reveal more specific effects. Additionally, it is unclear what constitutes a clinically significant change in the proportion of open to closed class words. This limits our ability to interpret the functional relevance of the change observed in our participants, especially in the context of scripted speech. Further investigation is warranted in order to determine whether proportions of open and closed class words are useful in characterizing treatment-induced change in connected speech in nfvPPA.

PERFORMANCE ON DISCOURSE MEASURES: UNTRAINED TOPICS

A significant reduction in the number of fluency disruptions per hundred words was found for untrained script topics from pre- to post-treatment. We did not expect to see generalization of treatment effects related to script training at the group level, given the lack of consistent evidence for transfer to untrained material in previous studies. This unexpected finding indicates that script training may have a generalized effect on the production of fillers, repetitions, revisions, and phonological fragments in connected speech. Given that the production of these structures may indicate difficulties with motor speech, syntax, or lexical retrieval, further research is required to determine the underlying mechanism for improvement.

This finding may suggest that script training results in generalized improvement in this dimension of fluency through repeated practice of an “over-articulation” strategy. It is also possible that participants simply gained a greater level of comfort with speaking by engaging in regular speech production practice, and relied less on corrective formulations and hesitations. The effect of repeated probing of untrained topics may have also played a role in improvement on this measure for untrained topics. However, the significant result for the interaction between time and training condition for this measure confirms the specificity of the training effect, regardless of the effects of repeating probing.

No other discourse measures reached significance for untrained topics but, as we predicted, some improvement was observed at the individual level. In a subset of participants, individual-level change scores for untrained topics showed numerical increases of equal or greater magnitude than the mean change for trained scripts. This was the case for WPM (n = 2), fluency disruptions per hundred words (n = 2), MLUm (n = 6), grammatical complexity (n = 5), and proportion of open to closed class words (n = 8). Taken together, these findings indicate that the greatest benefit of script training is observed for practiced material, but that individual participants may show generalized improvement in discourse production.

Additionally, the fact that outcome measures for untrained topics showed numerical increases for a subset of participants or were relatively stable from pre- to post-treatment may be meaningful in the context of a progressive syndrome, wherein the expected pattern of change is that of decline. In nonfluent stroke-induced aphasia, investigation of generalization of improvement following script training to untrained discourse generally has not been conducted in detail. Results from studies examining untrained connected speech relative to script training intervention provide evidence of generalization at the individual level on measures of grammaticality and informativeness, as well as rate of speech (Costello-Yacono & Balasubramanian, 2018; Nobis-Bosch et al., 2011). These outcomes indicate the need for further research assessing the potential for improvements related to script training in speech production and fluency to generalize to untrained connected speech in both nfvPPA and stroke-induced nonfluent aphasia. Future studies may find more in-depth analysis at the individual or sub-group level helpful in revealing more information regarding the effects of treatment on untrained connected speech tasks in heterogenous populations.

SPECIFICITY OF TREATMENT EFFECTS

Examination of the interaction of time (i.e., pre-treatment or post-treatment) and training status (i.e., trained or untrained) confirmed the specificity of the treatment effects. As anticipated, examination of the measure of correct, intelligible scripted words indicated a significant interaction such that trained scripts improved to a significantly greater degree from pre to post-treatment. The interaction term was also significant for WPM and number of fluency disruptions per hundred words. This, along with our initial analyses, indicate that trained topics improved to a significantly greater degree relative to untrained topics for these measures, which are generally related to speech production. With regard to measures indicative of grammaticality, a significant effect was observed for the interaction term for MLUm and proportion of open to closed class words. Although the fixed effect of time was only significant for trained topics, numerical improvements (MLUm and proportion of open to closed class words) were observed for untrained targets, explaining the significant effects. A significant effect was not found with regard to the interaction term and grammatical complexity. Given that trained topics were found to show significant improvement in grammatical complexity at post treatment, this may indicate that script training resulted in general improvement for both trained and untrained topics. This notion is supported by numerical increases observed for untrained topics on this measure. Further research with larger samples allowing for examination of subsets of participants is needed to explore the possibility for VISTA to generalize to grammatical measures. The interaction term did not trend towards significance for propositional density, indicating stability of performance from pre- to post-treatment for trained and untrained topics on these measures.

STRENGTHS, LIMITATIONS AND FUTURE DIRECTIONS

This is the largest study to examine the utility of discourse measures for characterizing treatment response in individuals with nfvPPA who have undergone script training. Findings from this study constitute a first step toward future efficacy research in which effect sizes can be taken into account to establish benchmarks for discourse-level change in the context of script training. These results may also inform power analyses which can be used to inform sample size and determine the robustness of the effects of script training.

Given the time- and labor-intensive nature of typical discourse analysis methods, an additional goal of this study was to investigate measures of grammatical ability and speech fluency that could be calculated with the greatest possible amount of automation, with ramifications for adoption in standard clinical settings. In this way, we hoped to discover means of quantifying treatment-related change in a large number of transcriptions without requiring extensive hand-coding or subjective judgements of grammaticality. Results confirmed that a number of largely automated measures (e.g., WPM, fluency disruptions per hundred words, MLUm, grammatical complexity, and proportion of open to closed class words) were sensitive to VISTA (for trained topics). However, more detailed analysis may be required in order to determine whether script training may indeed have generalized benefits for individuals with nfvPPA that weren't captured by largely automated analysis using CLAN.

Linguistic analysis procedures which require detailed hand-coding processes and determination of grammaticality or informativeness (e.g., analysis of correct information units) may be labor-intensive but robust options for future exploration of generalized treatment effects related to script training in nfvPPA. Several individuals demonstrated numerical improvements on measures related to grammaticality as well as speech production for untrained topics; therefore, future studies should explore the possibility for generalization effects in larger samples which would allow for analysis within subgroups that present with agrammatism or motor speech impairment to a greater degree. While the number of individuals in this study demonstrating minimal impairment of grammatical ability precluded analysis at the subgroup level, some numerical differences are apparent in the change scores. The group of five individuals with minimal grammatical impairment had a greater average change score (WPM M = 18.27; fluency disruptions per hundred words M = -16.72; MLUm M = 2.55; grammatical complexity M = 0.023; proportion of open to closed class words M = 0.116) than the average of the rest of the cohort (WPM M = 13.30; fluency disruptions per hundred words M = -12.43; MLUm M = 1.24; grammatical complexity M = 0.014; proportion of open to closed class words M = 0.035) for trained topics for all discourse measures excluding propositional density. However, it is unclear if this difference between the groups is significant in terms of response to script training treatment. Future research should explore whether script training provides differential benefit for individuals with relatively isolated motor speech impairment relative to those with relatively pure motor speech impairment or a mixed speech-language phenotype. Future studies should also explore cognitive-linguistic, motoric, and neural predictors of treatment response for specific discourse measures to better characterize individual participant factors that mediate treatment response.

Because our analyses focused on production of scripted content, future work should examine whether treatment-related improvements on relevant outcome measures generalize to less constrained connected speech tasks. Additionally, acoustic measures examining pauses and articulation rate (e.g., proportion of silence time, Vogel et al., 2017; formant centralization ratio, Sapir et al., 2010) should be employed in future research, as outcome measures in the current study were largely linguistic, providing a less precise characterization of motoric ability (i.e., via WPM and fluency disruptions).

CONCLUSION

Given the potential for discourse samples to approximate real-world communication abilities for persons with aphasia, discourse analysis holds potential as a complementary means of characterizing speech and language abilities and treatment-induced change. This study provides support for use of automatic analysis of transcribed connected speech in examining treatment effects for individuals with nfvPPA. Using minimal hand-coding of speech samples in conjunction with CLAN software, we derived discourse measures that captured speech-language production in a more nuanced manner. The automatic calculation of these measures, which were sensitive to treatment in this population, holds potential for application in standard clinical settings where time constraints preclude the regular use of discourse analysis in assessment and monitoring of treatment outcomes.

Our results provide further evidence that script training has the potential to address speech-language deficits for trained material in individuals with nfvPPA. Analysis of the connected speech of participants revealed improvements in measures of fluency, speech rate, and grammaticality for trained script topics. Given the functional nature of script training with personally-relevant topics, this intervention holds great promise in providing people with nfvPPA the means to continue communicating about meaningful subjects in their daily life.

Table 1: Demographics and Speech Language and Cognition Scores at Pre- and Post-Treatment

Patient ID	NFV1	NFV2	NFV3	NFV4	NFV5	NFV6	NFV7	NFV8	NFV9	NFV10	NFV11	NFV12	NFV13	NFV14	NFV15	NFV16	NFV17	NFV18	NFV19	NFV20	mean (SD)
Demographics																					
Age	69	71	57	68	72	61	71	69	75	64	61	68	60	77	78	71	65	68	75	69	68.45 (5.8)
Gender	male	female	male	female	female	male	female	female	female	male	female	female	male	female	female	female	male	male	female	male	-
Years of Education	18	15	18	12	16	18	14	16	16	13	20	20	17	22	16	14	14	20	16	18	16.65 (2.6)
Handedness	right	right	right	right	right	right	right	right	right	right	left	right	right	right	right	right	right	right	right	right	-
Speech, Language and Cognition																					
Mini-Mental State Examination PreTx (30)	26	29	28	30	28	25	24	25	29	24	28	30	28	25	23	25	30	30	29	30	27.3 (2.4)
Mini-Mental State Examination PostTx (30)	23	29	29	27	27	28	28	26	28	28	29	30	25	24	24	25	28	29	29	30	27.3 (2.1)
Boston Naming Test PreTx	55/60	56/60	58/60	53/60	54/60	56/60	23/30	52/60	57/60	21/30	30/30	30/30	30/30	12/30	28/30	28/30	30/30	29/30	30/30	30/30	-
Boston Naming Test PostTx	54/60	25/30	29/30	24/30	28/30	55/60	39/60	25/30	30/30	19/30	30/30	30/30	30/30	14/30	26/30	27/30	30/30	29/30	29/30	30/30	-
Western Aphasia Battery AQ PreTx (100)	80.6	85	96.9	92.6	85.8	78.3	80.2	83	84	76.2	91.4	97.2	96.6	65	78.8	79.7	87.4	96	96.8	96.8	86.42 (9.0)
Western Aphasia Battery AQ PostTx (100)	85.5	85.6	97.7	92.9	85.4	81.8	80	78.8	89	80.2	91.9	93.7	96.2	61.1	80	77.8	86.8	95.2	96.4	98	86.7 (9.1)
PPVT-short PreTx (16)	16	15	16	14	-	11	10	16	16	16	-	16	16	12	15	13	16	16	16	16	14.78 (2.0)
AOS rating PreTx[a] (0=none - 7=profound)	5	3	3	4	6	2	4	2	4	4	3	4	2	4	2	3	3	1	3	2	3.2 (1.2)
AOS rating PostTx[a] (0=none - 7=profound)	6	4	4	5	6	2	5	3	4	4	2	4	2	4	2	3	3	1	3	2	3.45 (1.4)
Dysarthria rating PreTx[a] (0=none - 7=profound)	4	0	4	4	5	2	3	0	4	3	1	0	0	0	3	1	1	1	1	0	1.85 (1.7)
Dysarthria rating PostTx[a] (0=none - 7=profound)	5	0	4	5	5	2	3	0	4	3	1	2	2	0	3	3	1	1	1	0	2.25 (1.7)
Northwestern Anagram Test (%) PreTx	40.0	53.3	90.0	90.0	50.0	-	36.7	66.7	90.0	56.7	46.7	73.3	93.3	63.3	60.0	16.7	66.7	60.0	100.0	70.0	64.39 (21.8)
Northwestern Anagram Test (%) PostTx	70.0	73.3	100.0	83.3	73.3	50.0	53.3	100.0	93.3	46.7	43.3	93.3	100.0	63.3	13.3	10.0	73.3	56.7	96.7	66.7	67.99 (26.8)

[a] from Wertz et al. (1984); AQ = Aphasia Quotient, PPVT = Peabody Picture Vocabulary Test, AOS = Apraxia of Speech

Table 2: Clinician-Guided VISTA Treatment Hierarchy

Participant is asked to:	
1.	Choose each correct script sentence from four foil sentences.
2.	Put the script sentences in the correct order.
3.	Read the entire script aloud.
4.	Produce individual scripted sentences in response to questions.
5.	Produce the entire script from memory.
6.	Respond to questions with scripted sentences out of the correct order of the
	During the second treatment session for a given script, a novel communication partner has an unscripted conversation with the participant to promote conversational usage of scripted material.

*Note. Feedback regarding articulation and grammar occurred during steps 3-6, with targeted practice, as needed.

Table 3: Pre- and Post-Treatment Performance on Trained and Untrained Topics

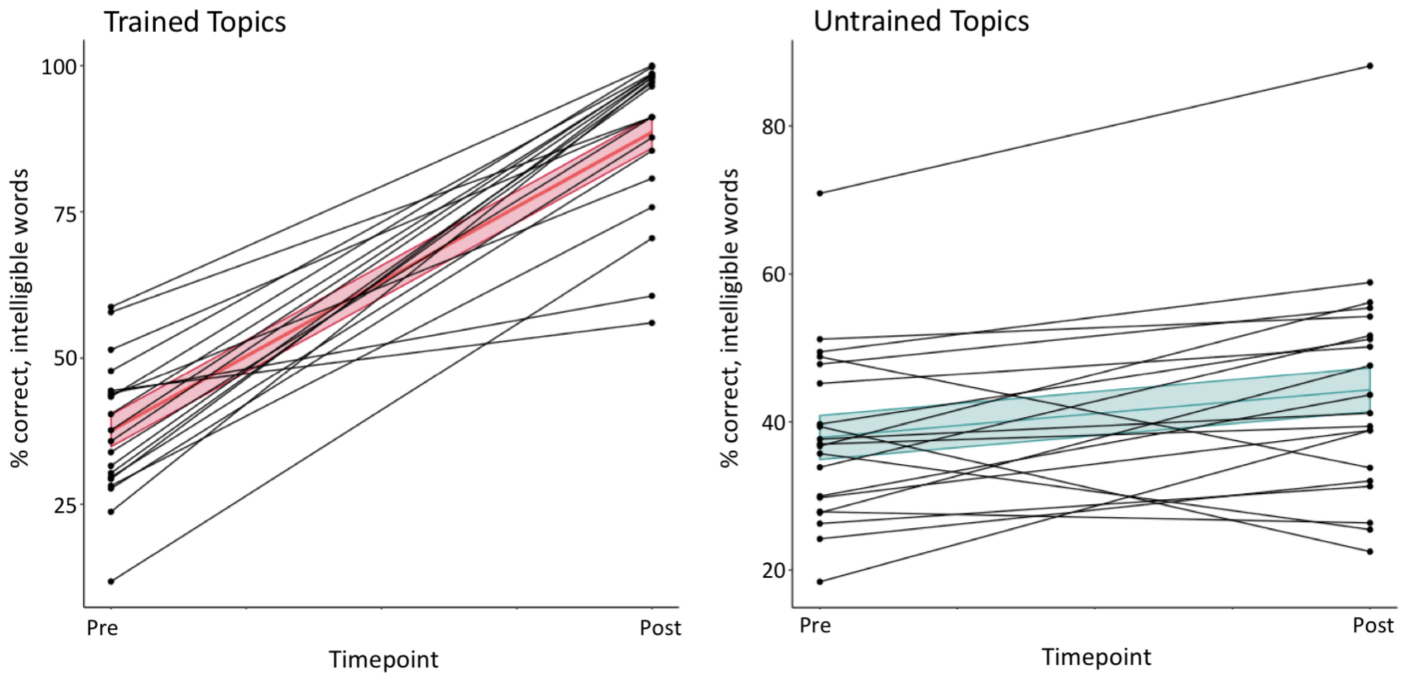
Timepoint	Primary Outcome Measure (percent correct intelligible scripted words)	MLUm (morphemes)	WPM	Fluency Disruptions/ 100 Words	Open Class/ Closed Class	Propositional Density	Grammatical Complexity
Trained Topics							
Pre Trained M (SD)	37.55 (11.70)	9.82 (3.49)	56.43 (20.44)	22.19 (13.35)	.98 (.36)	.46 (.08)	.07 (.04)
Post Trained M (SD)	88.63 (13.31)	11.39 (3.90)	71.18 (28.17)	8.69 (12.55)	1.07 (.34)	.47 (.06)	.09 (.04)
Untrained Topics							
Timepoint	**Primary Outcome Measure** (percent correct intelligible scripted words)	**MLUm** (morphemes)	**WPM**	**Fluency Disruptions/ 100 Words**	**Open Class/ Closed Class**	**Propositional Density**	**Grammatical Complexity**
Pre Untrained M (SD)	37.89 (11.99)	10.00 (3.69)	55.48 (21.78)	22.17 (13.09)	1.02 (.43)	.46 (.06)	.07 (.03)
Post Untrained M (SD)	44.35 (15.07)	10.52 (3.76)	59.25 (23.75)	11.05 (11.14)	.97 (.31)	.45 (.06)	.07(.03)

*Note. M= Mean, SD= standard deviation

Table 4: Change Scores Pre- to Post-Treatment for Trained and Untrained Topics

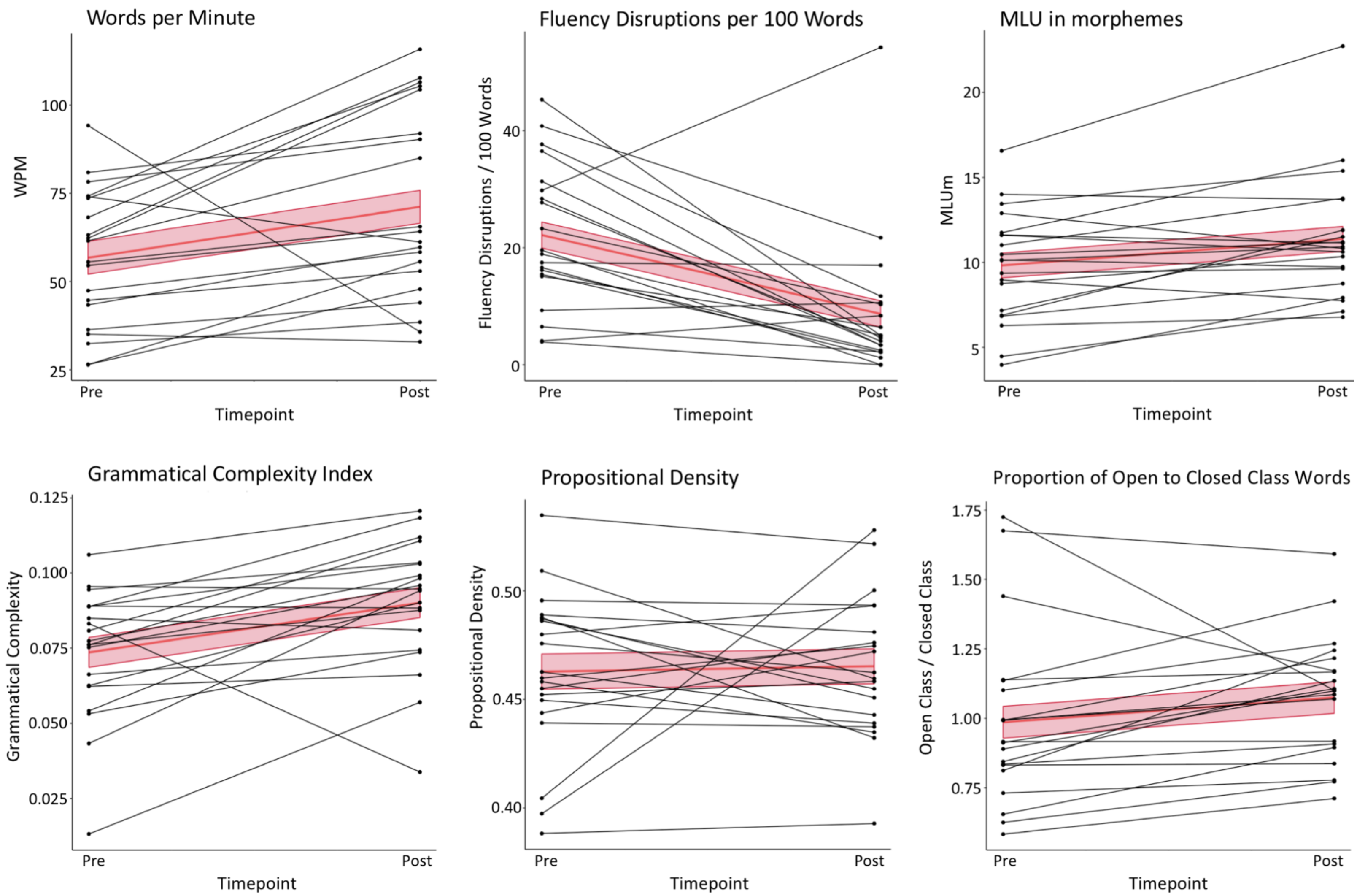
Participant	WPM		Fluency Disruptions/ 100 Words		MLUm (morphemes)		Grammatical Complexity		Propositional Density		Open Class Words/ Closed Class Words	
	Trained	Untrained	Trained	Untrained	Trained	Untrained	Trained	Untrained	Trained	Untrained	Trained	Untrained
NFV1	3.95	3.05	29.20	4.24	-13.91	5.19	-0.084	-1.190	0.005	-0.230	0.044	-0.020
NFV2	5.01	-1.17	21.37	5.42	-40.30	-11.43	-0.623	-0.387	0.103	0.009	0.044	-0.011
NFV3	6.15	2.93	11.06	12.02	-28.02	-14.45	0.287	-0.108	-0.036	-0.012	-0.001	-0.005
NFV4	4.33	1.09	9.65	5.99	-13.69	-11.78	0.075	0.018	-0.013	-0.017	0.051	0.008
NFV5	1.60	0.86	8.27	5.72	-13.28	-11.55	0.290	0.108	-0.050	-0.033	0.028	0.038
NFV6	-1.00	-1.86	41.65	-18.21	-23.67	21.45	0.432	0.115	-0.056	-0.016	-0.001	-0.015
NFV7	0.49	0.83	31.76	18.97	-19.60	-0.29	0.146	-0.319	0.028	0.049	0.011	0.013
NFV8	0.27	1.83	-58.49	-14.34	24.46	3.55	0.241	-0.020	-0.008	0.000	0.009	-0.008
NFV9	0.81	0.94	7.69	2.04	-12.95	-10.97	0.005	-0.121	0.021	0.092	0.022	-0.013
NFV10	1.92	1.01	16.35	-16.85	-25.91	5.79	0.209	-0.009	-0.002	-0.034	0.020	0.003
NFV11	-2.07	-2.93	12.04	20.67	1.31	-5.96	0.224	-0.024	-0.031	-0.022	-0.004	0.028
NFV12	2.75	1.12	10.02	-3.26	-4.32	-3.55	0.073	0.115	-0.002	-0.036	0.014	-0.011
NFV13	1.94	2.02	23.45	10.23	-25.02	-22.26	0.047	0.000	-0.013	-0.003	0.031	0.036
NFV14	0.73	1.02	-12.90	-21.54	-0.45	0.43	0.129	0.209	0.013	-0.004	0.030	0.007
NFV15	-1.22	0.16	-2.21	0.50	-8.70	1.18	0.194	-0.090	-0.023	0.013	-0.049	0.013
NFV16	2.64	-0.59	10.95	7.08	4.29	6.08	-0.834	-0.259	0.123	0.023	0.008	-0.006
NFV17	-0.42	1.64	6.09	2.60	-19.06	-6.71	0.001	0.340	0.015	-0.049	0.004	0.043
NFV18	4.25	3.73	43.30	5.43	-31.93	-5.49	0.093	0.431	-0.020	0.011	0.021	0.024
NFV19	-0.47	0.05	39.51	21.70	-3.91	-4.20	0.168	0.114	-0.011	0.021	0.015	-0.013
NFV20	-0.31	-5.24	42.07	13.35	-15.40	-4.95	0.029	0.095	0.006	0.043	0.035	0.003

Figure 1: Fixed Effect of Time on Script Accuracy



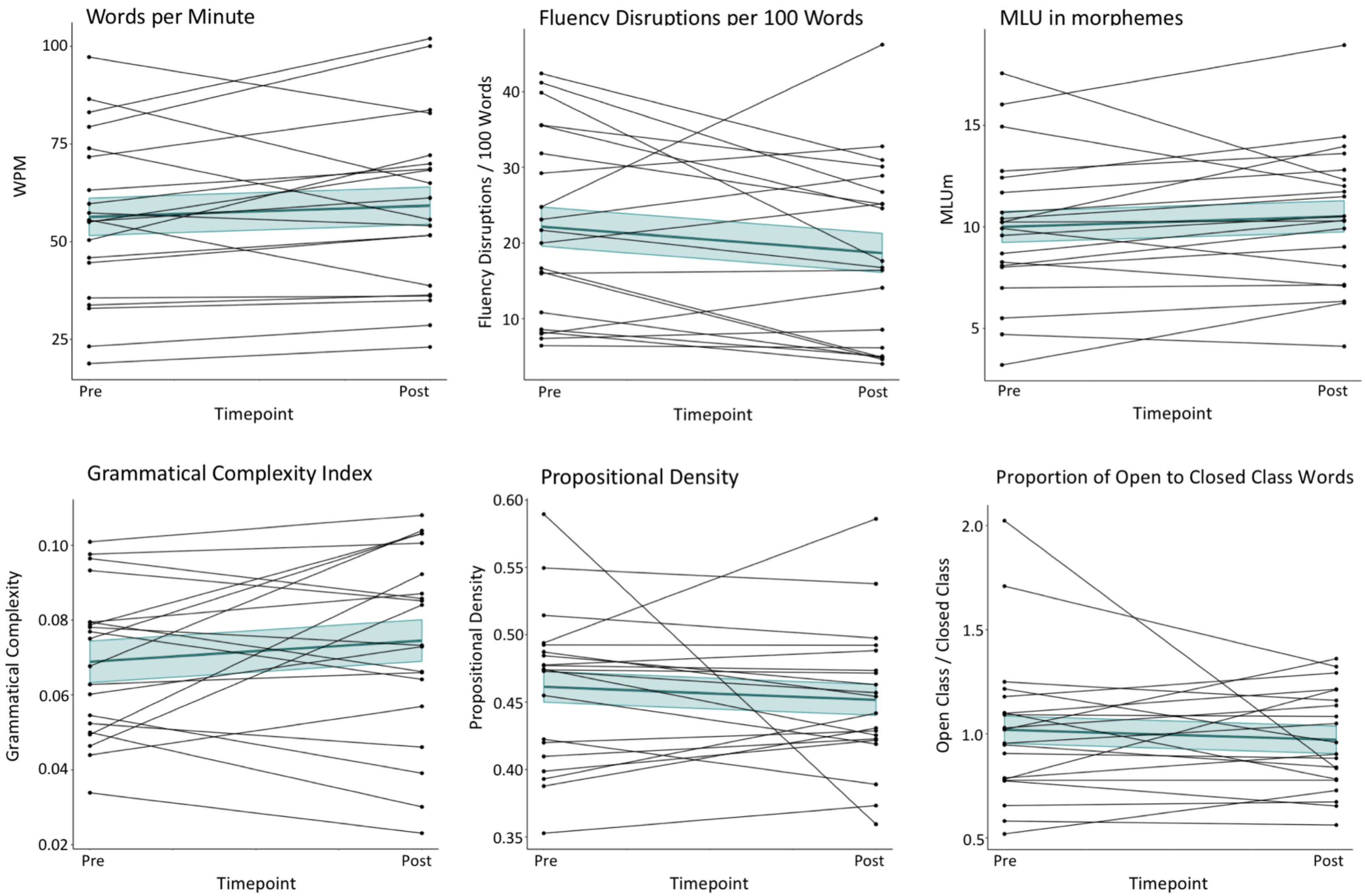
Note. Each model includes a random intercept for participant. The y axis presents fitted values from the linear mixed effects model. The fitted regression line and standard error are shown in color. Each participant's average performance across scripts at pre- and post-treatment are shown in black.

Figure 2: Fixed Effect of Time on Discourse Measures for Trained Topics



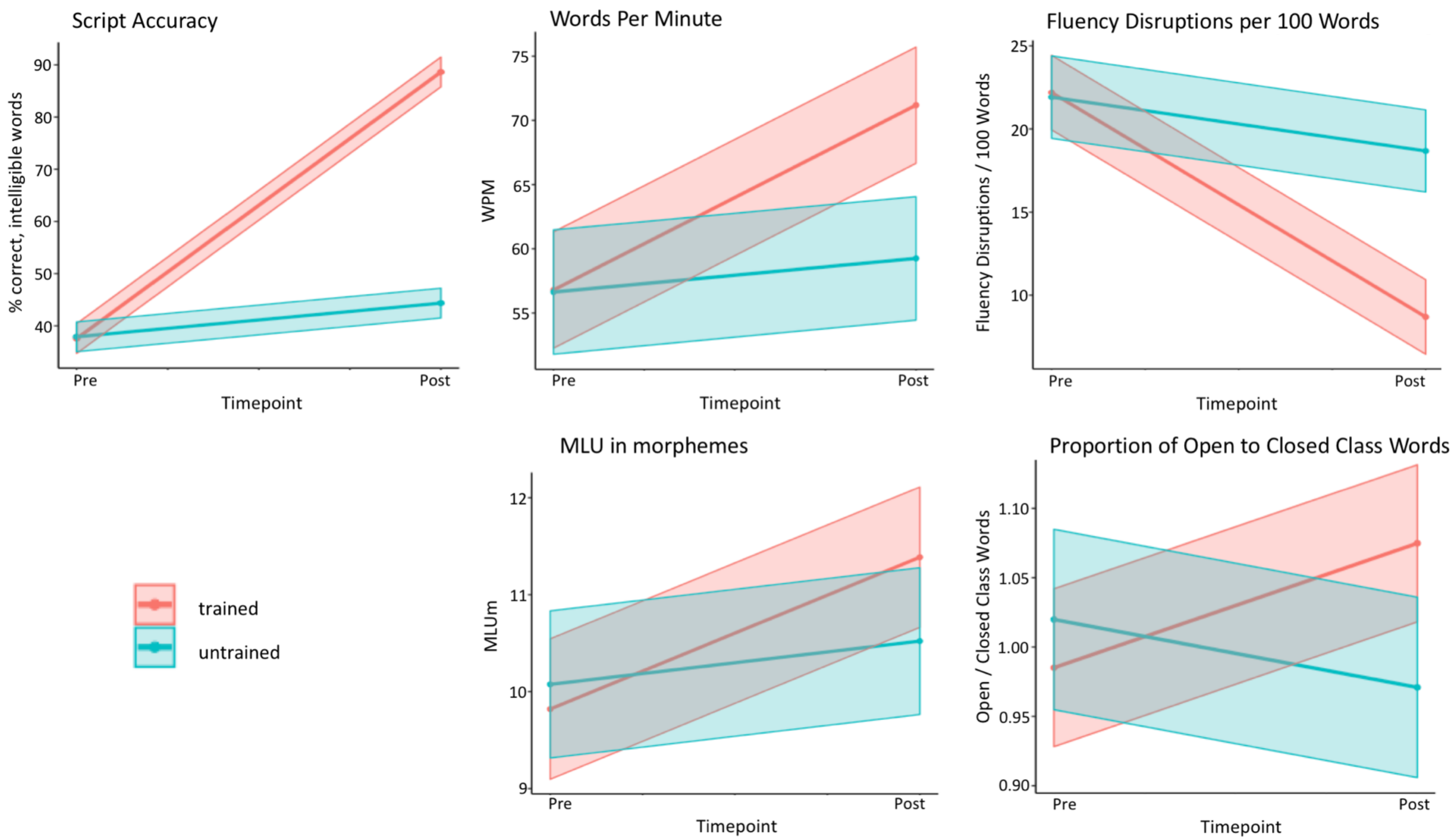
Note. Each model includes a random intercept for participant. The y axis presents fitted values from the linear mixed effects model. The fitted regression line and standard error are shown in color. Each participant's average performance across scripts at pre- and post-treatment are shown in black.

Figure 3: Fixed Effect of Time on Discourse Measures for Untrained Topics



Note. Each model includes a random intercept for participant. The y axis presents fitted values from the linear mixed effects model. The fitted regression line and standard error are shown in color. Each participant's average performance across scripts at pre- and post-treatment are shown in black.

Figure 4: Significant Interactions of Time and Training Condition



Note. Each model includes a random intercept for participant. The y axis presents fitted values from the linear mixed effects model. Standard error is shown in shaded color along the fitted regression line.

References

Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology*, *29*(2), 130–131.

Ali, N., Rafi, M. S., Khan, M. S. G., & Mahfooz, U. (2018). The effectiveness of script training to restore lost communication in a patient with Broca's aphasia. *Journal of the Pakistan Medical Association*, *68*(7), 1070–1075.

Ash, S., Evans, E., Shea, J. O., Powers, J., Boller, A., Weinberg, D., Haley, J., Mcmillan, C., Irwin, D. J., Rascovsky, K., & Grossman, M. (2013). *Differentiating primary progressive aphasias in a brief sample of connected speech*. 329–336.

Ash, S., McMillan, C., Gunawardena, D., Avants, B., Morgan, B., Khan, A., Moore, P., Gee, J., & Grossman, M. (2010). Speech errors in progressive non-fluent aphasia. *Brain and Language*, *113*(1), 13–20. https://doi.org/10.1016/j.bandl.2009.12.001

Ash, S., Moore, P., Antani, S., McCawley, G., Work, M., & Grossman, M. (2006). Trying to tell a tale: Discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology*, *66*(9), 1405–1413. https://doi.org/10.1212/01.wnl.0000210435.72614.38

Ash, S., Moore, P., Vesely, L., Gunawardena, D., McMillan, C., Anderson, C., Avants, B., & Grossman, M. (2009). Non-fluent speech in frontotemporal lobar degeneration. *Journal of Neurolinguistics*, *22*(4), 370–383. https://doi.org/10.1016/j.jneuroling.2008.12.001

Ash, S., Nevler, N., Phillips, J., Irwin, D. J., McMillan, C. T., Rascovsky, K., & Grossman, M. (2019). A longitudinal study of speech production in primary progressive aphasia and behavioral variant frontotemporal dementia. *Brain and*

Language, *194*(February), 46–57. https://doi.org/10.1016/j.bandl.2019.04.006

Austermann, S. N., Robin, D. A., Maas, E., Ballard, K. J., & Schmidt, R. A. (2008). Effects of Feedback Frequency and Timing on Acquisition, Retention, and Transfer of Speech Skills in Acquired Apraxia of Speech. *Journal of Speech, Language, and Hearing Research*, *51*(October), 1088–1113. https://doi.org/10.1044/1092-4388(2008/06-0042)

Ballard, K. J., Maas, E., & Robin, D. A. (2007). Treating control of voicing in apraxia of speech with variable practice. *Aphasiology*, *21*(12), 1195–1217. https://doi.org/10.1080/02687030601047858

Beber, B. C., Berbert, M. C. B., Grawer, R. S., & Cardoso, M. C. de A. F. (2018). Estratégias de controle de ritmo e taxa de fala para a apraxia de fala na afasia progressiva primária não-fluente. *Dementia e Neuropsychologia*, *12*(1), 80–84. https://doi.org/10.1590/1980-57642018dn12-010012

Beeke, S., Maxim, J., & Wilkinson, R. A. Y. (2008). *Rethinking agrammatism : Factors affecting the form of language elicited via clinical test procedures*. *22*(May), 317–323. https://doi.org/10.1080/02699200801918911

Bilda, K. (2011). Video-based conversational script training for aphasia: A therapy study. *Aphasiology*, *25*(2), 191–201. https://doi.org/10.1080/02687031003798254

Bird, H., & Franklin, S. (1996). Cinderella revisited: A comparison of fluent and non-fluent aphasic speech. *Journal of Neurolinguistics*, *9*(3), 187–206. https://doi.org/10.1016/0911-6044(96)00006-1

Bonilha, L., Hillis, A. E., Wilmskoetter, J., Hickok, G., Basilakos, A., Munsell, B., Rorden, C., & Fridriksson, J. (2019). Neural structures supporting spontaneous and assisted (entrained) speech fluency. *Brain*, *142*(12), 3951–3962. https://doi.org/10.1093/brain/awz309

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of prepositional idea density from part-of-speech tagging. *Behavior Research Methods*, *40*(2), 540–545. https://doi.org/10.3758/BRM.40.2.540

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics and Phonetics*, *30*(7), 489–518. https://doi.org/10.3109/02699206.2016.1145740

Bryant, L., Spencer, E., Ferguson, A., Craig, H., Colyvas, K., & Worrall, L. (2013). Propositional Idea Density in aphasic discourse. *Aphasiology*, *27*(8), 992–1009. https://doi.org/10.1080/02687038.2013.803514

Butler, S. M., & Snowdon, D. A. (1996). Trends in mortality in older women: Findings from the Nun Study. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *51*(4), 201–208. https://doi.org/10.1093/geronb/51B.4.S201

Cadório, I., Lousada, M., Martins, P., & Figueiredo, D. (2017). Generalization and maintenance of treatment gains in primary progressive aphasia (PPA): a systematic review. *International Journal of Language and Communication Disorders*, *52*(5), 543–560. https://doi.org/10.1111/1460-6984.12310

Cherney, L. R., Braun, E. J., Lee, J. B., Kocherginsky, M., & Van Vuuren, S. (2019). Optimising recovery in aphasia: Learning following exposure to a single dose of computer-based script training. *International Journal of Speech-Language Pathology*, *21*(5), 448–458. https://doi.org/10.1080/17549507.2019.1661518

Cherney, L. R., & Halper, A. S. (2008). Novel technology for treating individuals with aphasia and concomitant cognitive deficits. *Topics in Stroke Rehabilitation*, *15*(6), 542–554. https://doi.org/10.1310/tsr1506-542

Cherney, L. R., Halper, A. S., Holland, A. L., & Cole, R. (2008). Computerized script training for aphasia: Preliminary results. *American Journal of Speech-Language Pathology*, *17*(1), 19–34. https://doi.org/10.1044/1058-0360(2008/003)

Cherney, L. R., Halper, A. S., & Kaye, R. C. (2011). Computer-based script training for aphasia: Emerging themes from post-treatment interviews. *Journal of Communication Disorders*, *44*(4), 493–501. https://doi.org/10.1016/j.jcomdis.2011.04.002

Cherney, L. R., Kaye, R. C., Lee, J. B., & van Vuuren, S. (2015). Impact of personal relevance on acquisition and generalization of script training for aphasia: a preliminary analysis. *Journal of Speech, Language, and Hearing Research*, *24*(Supplement: Select Papers from the 44th Clinical Aphasiology Conference), S913–S922.

Cherney, L. R., Kaye, R. C., & Van Vuuren, S. (2014). Acquisition and maintenance of scripts in aphasia: a comparison of two cuing conditions. *American Journal of Speech-Language Pathology*, *23*(2), S343–S360. https://doi.org/10.1044/2014_AJSLP-13-0097

Costello-Yacono, M., & Balasubramanian, V. (2018). A Comparison of Two Treatment Approaches for Agrammatic Broca's Aphasia: Script Therapy vs. Verb Network Strengthening Treatment. *Frontiers in Human Neuroscience*, *12*. https://doi.org/10.3389/conf.fnhum.2018.228.00020

Croot, K., Ballard, K., Leyton, C. E., & Hodges, J. R. (2012). Apraxia of speech and phonological errors in the diagnosis of nonfluent/agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1562–1572. https://doi.org/10.1044/1092-4388(2012/11-0323)

Croot, K., Nickels, L., Laurence, F., & Manning, M. (2009). Impairment- and activity/participation-directed interventions in progressive language impairment: Clinical and theoretical issues. *Aphasiology*, *23*(2), 125–160. https://doi.org/10.1080/02687030801943179

Crystal, D., Fletcher, P., & Garman, M. (1976). *The grammatical analysis of language disability: A procedure for assessment and remediation* (Vol. 1). Edward Arnold.

Dickerson, B. C. (2011). Quantitating severity and progression in primary progressive aphasia. *Journal of Molecular Neuroscience*, *45*(3), 618–628. https://doi.org/10.1007/s12031-011-9534-2

Dronkers, N. F. (1996). A new brain region for coordinating speech production. *Nature*, *384*(November), 14. https://www.nature.com/articles/384159a0.pdf

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test (3rd ed.)*. American Guidance Service.

Edmonds, L., Nadeau, S., Kiran, S. (2009). Effect of VNeST on Lexical Retrieval of Content Words in Sentences in Persons with Aphasia. *Aphasiology*, *23*(3), 402–424. https://doi.org/10.1080/02687030802291339.Effect

Ferguson, A., Spencer, E., Craig, H., & Colyvas, K. (2013). Propositional Idea Density in women's written language over the lifespan: Computerized analysis. *Cortex*, *55*(1), 107–121. https://doi.org/10.1016/j.cortex.2013.05.012

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *23*, 221–

223.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*, *55*(1), 43–60. https://doi.org/10.1016/j.cortex.2012.12.006

Fridriksson, J., Basilakos, A., Hickok, G., Bonilha, L., & Rorden, C. (2015). Speech entrainment compensates for Broca's area damage. *Cortex*, *69*, 68–75. https://doi.org/10.1016/j.cortex.2015.04.013

Fridriksson, J., Hubbard, H. I., Hudspeth, S. G., Holland, A. L., Bonilha, L., Fromm, D., & Rorden, C. (2012). Speech entrainment enables patients with Broca's aphasia to produce fluent speech. *Brain*, *135*(12), 3815–3829. https://doi.org/10.1093/brain/aws301

Fromm, D., Greenhouse, J., Hou, K., Russell, G. A., Cai, X., Forbes, M., Holland, A., & MacWhinney, B. (2016). Automated proposition density analysis for discourse in aphasia. *Journal of Speech, Language, and Hearing Research*, *59*(October), 1123–1132.

Goldberg, S., Haley, K. L., & Jacks, A. (2012). Script training and generalization for people with aphasia. *American Journal of Speech-Language Pathology*, *21*(3), 222–238. https://doi.org/10.1044/1058-0360(2012/11-0056)

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W., & Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, *55*(3), 335–346.

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Graham, N. L., Leonard, C., Tang-Wai, D. F., Black, S., Chow, T. W., Scott, C. J. M., McNeely, A. A., Masellis, M., & Rochon, E. (2016). Lack of Frank Agrammatism in the Nonfluent Agrammatic Variant of Primary Progressive Aphasia. *Dementia and Geriatric Cognitive Disorders Extra*, *6*(3), 407–423. https://doi.org/10.1159/000448944

Graham, N. L., Patterson, K., & Hodges, J. R. (2004). When more yields less: Speaking and writing deficits in nonfluent progressive aphasia. *Neurocase*, *10*(2), 141–155. https://doi.org/10.1080/13554790409609945

Grande, M., Hussmann, K., Bay, E., Christoph, S., Piefke, M., Willmes, K., & Huber, W. (2008). Basic parameters of spontaneous speech as a sensitive method for measuring change during the course of aphasia. *International Journal of Language and Communication Disorders*, *43*(4), 408–426. https://doi.org/10.1080/13682820701685991

Grasso, S. M., Cruz, D. F., Benavidez, R., Peña, E. D., & Henry, M. L. (2019). Video-implemented script training in a bilingual spanish–english speaker with aphasia. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2295–2316. https://doi.org/10.1044/2018_JSLHR-L-18-0048

Grossman, M. (2012). The non-fluent/agrammatic variant of primary progressive aphasia. *The Lancet Neurology*, *11*(6), 545–555. https://doi.org/10.1016/S1474-4422(12)70099-6

Grossman, M., Mickanin, J., Onishi, K., Hughes, E., D'Esposito, M., Ding, X. S., Alavi, A., & Reivich, M. (1996). Progressive nonfluent aphasia: Language, cognitive, and PET measures contrasted with probable Alzheimer's disease. *Journal of Cognitive Neuroscience*, *8*(2), 135–154. https://doi.org/10.1162/jocn.1996.8.2.135

Hameister, I., Nickels, L., Abel, S., & Croot, K. (2017). "Do you have mowing the lawn?"–improvements in word retrieval and grammar following constraint-induced language therapy in primary progressive aphasia. *Aphasiology*, *31*(3), 308–331. https://doi.org/10.1080/02687038.2016.1197558

Harciarek, M., Sitek, E. J., & Kertesz, A. (2014). The patterns of progression in primary progressive aphasia-Implications for assessment and management. *Aphasiology*, *28*(8–9), 964–980. https://doi.org/10.1080/02687038.2014.904498

Helm-Estabrooks, N., Nicholas, M., & Helm, S. A. (2000). *Sentence production program for aphasia*. Pro-ed.

Henry, M. L., Hubbard, H. I., Grasso, S. M., Mandelli, M. L., Wilson, S. M., Sathishkumar, M. T., Fridriksson, J., Daigle, W., Boxer, A. L., Miller, B. L., & Gorno-Tempini, M. L. (2018). Retraining speech production and fluency in non-fluent/agrammatic primary progressive aphasia. *Brain*, *141*(6), 1799–1814.

https://doi.org/10.1093/brain/awy101

Henry, M. L., Meese, M. V., Truong, S., Babiak, M. C., Miller, B. L., & Gorno-Tempini, M. L. (2013). Treatment for apraxia of speech in nonfluent variant primary progressive aphasia. *Behavioural Neurology*, *26*(1–2), 77–88. https://doi.org/10.3233/BEN-2012-120260

Herbert, R., Hickin, J., Howard, D., & Best, W. (2008). *Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia ? 22*(2), 184–203. https://doi.org/10.1080/02687030701262613

Holland, A., Munoz, M., Milman, L., & Bays, G. (2002). Scripts, scenarios, and treatment of aphasia. *Paper Presented to the World Federation of Neurology, Aphasia, & Cognitive Disorders Study Group*.

Hsu, C. J., & Thompson, C. K. (2018). Manual versus automated narrative analysis of agrammatic production patterns: The northwestern narrative language analysis and computerized language analysis. *Journal of Speech, Language, and Hearing Research*, *61*(2), 373–385. https://doi.org/10.1044/2017_JSLHR-L-17-0185

Jokel, R., Graham, N. L., Rochon, E., & Leonard, C. (2014). Word retrieval therapies in primary progressive aphasia. *Aphasiology*, *28*(8–9), 1038–1068.

Josephs, K. A., Duffy, J. R., Strand, E. A., MacHulda, M. M., Senjem, M. L., Master, A. V., Lowe, V. J., Jack, C. R., & Whitwell, J. L. (2012). Characterizing a neurodegenerative syndrome: Primary progressive apraxia of speech. *Brain*, *135*(5), 1522–1536. https://doi.org/10.1093/brain/aws032

Jung, Y., Duffy, J. R., & Josephs, K. A. (2013). Primary progressive aphasia and apraxia of speech. *Seminars in Neurology*, *33*(4), 342–347. https://doi.org/10.1055/s-0033-

1359317

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*. Lippincott, Williams & Wilkins.

Kemper, S., & Kemper, S. (2006). The cost of doing two things at once. *Aging, Neuropsychology, and Cognition*, *13*, 115–139.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity adn propositional content. *Psychology and Aging*, *16*, 600–614. https://doi.org/10.1037/0882-7974.16.2.312

Kertesz, A. (2006). *Western Aphasia Battery -Revised (WAB-R)*.

Knibb, J. A., Woollams, A. M., Hodges, J. R., & Patterson, K. (2009). Making sense of progressive non-fluent aphasia: An analysis of conversational speech. *Brain*, *132*(10), 2734–2746. https://doi.org/10.1093/brain/awp207

Kortte, K. B., & Rogalski, E. J. (2013). Behavioural interventions for enhancing life participation in behavioural variant frontotemporal dementia and primary progressive aphasia. *International Review of Psychiatry*, *25*(2), 237–245. https://doi.org/10.3109/09540261.2012.751017

Lee, J. B., Kaye, R. C., & Cherney, L. R. (2009). Conversational script performance in adults with non-fluent aphasia: Treatment intensity and aphasia severity. *Aphasiology*, *23*(7–8), 885–897. https://doi.org/10.1080/02687030802669534

Logan, G. D. (1988). Toward an Instance Theory of Automatization. *Psychological Review*, *95*(4), 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Lum, C., & Ellis, A. W. (1999). Why do some aphasics show an advantage on some tests of nonpropositional (automatic) speech? *Brain and Language*, *70*(1), 95–118. https://doi.org/10.1006/brln.1999.2147

Machado, T. H., Campanha, A. C., Caramelli, P., & Carthery-Goulart, M. T. (2014). Brief intervention for agrammatism in Primary Progressive Nonfluent Aphasia: A case report. *Dementia & Neuropsychologia*, *8*(3), 291–296. https://doi.org/10.1590/s1980-57642014dn83000014

Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M. M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? *Neuropsychologia*, *77*, 211–222. https://doi.org/10.1016/j.neuropsychologia.2015.08.019

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum Associates.

Manheim, L. M., Halper, A. S., & Cherney, L. (2009). Patient-Reported Changes in Communication After Computer-Based Script Training for Aphasia. *Archives of Physical Medicine and Rehabilitation*, *90*(4), 623–627. https://doi.org/10.1016/j.apmr.2008.10.022

Mayer, J. F., & Murray, L. L. (2003). *Functional measures of naming in aphasia : Word retrieval in confrontation naming versus connected speech*. *17*(5). https://doi.org/10.1080/02687030344000148

Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, *11*(6). https://doi.org/10.1002/ana.410110607

Mesulam, M. M., Wieneke, C., Rogalski, E., Cobia, D., Thompson, C., & Weintraub, S. (2009). Quantitative template for subtyping primary progressive aphasia. *Archives of*

Neurology, *66*(12), 1545–1551. https://doi.org/10.1001/archneurol.2009.288

Meyer, A. M., Tippett, D. C., Turner, R. S., & Friedman, R. B. (2019). Long-term maintenance of anomia treatment effects in primary progressive aphasia. *Neuropsychological Rehabilitation*, *29*(9), 1439–1463. https://doi.org/10.1080/09602011.2018.1425146

Miller, J., & Iglesias, A. (2012). *Systematic analysis of language transcripts (SALT)* (Research Version 2012). SALT Software LLC.

Montembeault, M., Brambati, S. M., Gorno-Tempini, M. L., & Migliaccio, R. (2018). Clinical, anatomical, and pathological features in the three variants of primary progressive aphasia: A review. *Frontiers in Neurology*, *9*(AUG). https://doi.org/10.3389/fneur.2018.00692

Mortimer, J., & Borenstein, A. (2012). Idea density in early life is associated with APOE-ε4 but affects the risk of Alzheimer's disease independently. *Alzheimer's & Dementia*, *8*, P629–P629. https://doi.org/10.1016/j.jalz.2012.05.1677

Moss, L. (2009). Functional discourse and quality of life changes following conversational training in a group setting. *Presentation at: Clinical Aphasiology Conference*.

Murray, L. L. (1998). Longitudinal treatment of primary progressive aphasia: a case study. *Aphasiology*, *12*(7), 651–672.

Nicholas, L. E., & Brookshire, R. H. (1993). *A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia*. *36*(April), 338–350.

Nobis-Bosch, R., Springer, L., Radermacher, I., & Huber, W. (2011). Supervised home training of dialogue skills in chronic aphasia: A randomized parallel group study. *Journal of Speech, Language, and Hearing Research*, *54*, 1118–1136. https://doi.org/10.1044/1092-4388(2010/09-0204)

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). *nlme: Linear and Nonlinear Mixed Effects Models* (R package version 3.1-145). https://cran.r-project.org/package=nlme

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Riley, K. P., Snowdon, D. A., Desrosiers, M. F., & Markesbery, W. R. (2005). Early life linguistic ability, late life cognitive function, and neuropathology: Findings from the Nun Study. *Neurobiology of Aging*, *26*, 341–347. https://doi.org/10.1016/j.neurobiolaging.2004.06.019

Rising, K. (2014). Treatment for lexical retrieval in progressive aphasia. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *24*(4), 137–144.

Rochon, E., Laird, L., Bose, A., & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation*, *15*(1), 1–36. https://doi.org/10.1080/09602010343000327

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, *72*(3), 193–218. https://doi.org/10.1006/brln.1999.2285

Rogalski, E. J., Cobia, D., Harrison, T. M., Wieneke, C., Thompson, C. K., Weintraub, S., & Mesulam, M. M. (2011). Anatomy of language impairments in primary progressive aphasia. *Journal of Neuroscience*, *31*(9), 3344–3350.

https://doi.org/10.1523/JNEUROSCI.5544-10.2011

Rogalski, E. J., & Mesulam, M. M. (2009). Clinical trajectories and biological features of primary progressive aphasia (PPA). *Current Alzheimer Research*, *6*(4), 331–336.

Rohrer, J. D., Rossor, M. N., & Warren, J. D. (2010a). Apraxia in progressive nonfluent aphasia. *Journal of Neurology*, *257*(4), 569–574. https://doi.org/10.1007/s00415-009-5371-4

Rohrer, J. D., Rossor, M. N., & Warren, J. D. (2010b). Syndromes of nonfluent primary progressive aphasia: A clinical and neurolinguistic analysis. *Neurology*, *75*(7), 603–610. https://doi.org/10.1212/WNL.0b013e3181ed9c6b

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*(3), 440–479. https://doi.org/10.1016/0093-934X(89)90030-8

Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012). Abnormalities of connected speech in the non-semantic variants of primary progressive aphasia. *Aphasiology*, *26*(10), 1219–1237. https://doi.org/10.1080/02687038.2012.710318

Santos-Santos, M. A., Mandelli, M. L., Binney, R. J., Ogar, J., Wilson, S. M., Henry, M. L., Hubbard, H. I., Meese, M., Attygalle, S., Rosenberg, L., Pakvasa, M., Trojanowski, J. Q., Grinberg, L. T., Rosen, H., Boxer, A. L., Miller, B. L., Seeley, W. W., & Gorno-Tempini, M. L. (2016). Features of patients with nonfluent/agrammatic primary progressive aphasia with underlying progressive supranuclear palsy pathology or corticobasal degeneration. *JAMA Neurology*, *73*(6), 733–742. https://doi.org/10.1001/jamaneurol.2016.0412

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant centralization ratio:

A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research*, *53*(1), 114–125. https://doi.org/10.1044/1092-4388(2009/08-0184)

Schneider, S. L., Thompson, C. K., & Luring, B. (1996). Effects of verbal plus gestural matrix training on sentence production in a patient with primary progressive aphasia. *Aphasiology*, *10*(3), 297–317. https://doi.org/10.1080/02687039608248414

Shewan, M. (1988). Expressive language recovery in aphasia using the Shewan Spontaneous Language Analysis (SSLA) system. *Journal of Communication Disorders*, *21*, 155–169.

Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. *JAMA*, *275*, 528–532.

Spencer, E., Ferguson, A., Craig, H., Colyvas, K., Hankey, G. J., & Flicker, L. (2015). Propositional idea density in older men's written language: Findings from the HIMS study using computerised analysis. *Clinical Linguistics and Phonetics*, *29*(2), 85–101. https://doi.org/10.3109/02699206.2014.956263

Spinelli, E. G., Mandelli, M. L., Miller, Z. A., Miguel, A., Henry, M. L., Comi, G., Rabinovici, G., & Rosen, H. J. (2017). *Typical and atypical pathology in primary progressive aphasia variants*. *81*(3), 430–443.

Szabo, G. B., Fromm, D., Heimlich, T., & Holland, A. (2014). Script training and its application to everyday life observed in an aphasia center. *Paper Presented at the Clinical Aphasiology Conference; May 2014; St. Simons Island, GA*.

Tetzloff, K. A., Duffy, J. R., Clark, H. M., Utianski, R. L., Strand, E. A., Machulda, M.

M., Botha, H., Martin, P. R., Schwarz, C. G., Senjem, M. L., Reid, R. I., Gunter, J. L., Spychalla, A. J., Knopman, D. S., Petersen, R. C., Jack, C. R., Lowe, V. J., Josephs, K. A., & Whitwell, J. L. (2019). Progressive agrammatic aphasia without apraxia of speech as a distinct syndrome. *Brain : A Journal of Neurology*, *142*(8), 2466–2482. https://doi.org/10.1093/brain/awz157

Tetzloff, K. A., Utianski, R. L., Duffy, J. R., Clark, H. M., Strand, E. A., Josephs, K. A., & Whitwell, J. L. (2018). Quantitative analysis of agrammatism in agrammatic primary progressive aphasia and dominant apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2337–2346. https://doi.org/10.1044/2018_JSLHR-L-17-0474

Thompson, C. K. (2013). *Northwestern narrative language analysis (NNLA) theory and methodology*.

Thompson, C. K., Ballard, K. J., Tait, M. E., Weintraub, S., & Mesulam, M. (1997). Patterns of language decline in non-fluent primary progressive aphasia. *Aphasiology*, *11*(4–5), 297–321. https://doi.org/10.1080/02687039708248473

Thompson, C. K., Cho, S., Hsu, C. J., Wieneke, C., Rademaker, A., Weitner, B. B., Mesulam, M. M., & Weintraub, S. (2012). Dissociations between fluency and agrammatism in primary progressive aphasia. *Aphasiology*, *26*(1), 20–43. https://doi.org/10.1080/02687038.2011.584691

Thompson, C. K., Meltzer-Asscher, A., Cho, S., Lee, J., Wieneke, C., Weintraub, S., & Mesulam, M. M. (2013). Syntactic and morphosyntactic processing in stroke-induced and primary progressive aphasia. *Behavioural Neurology*, *26*(1–2), 35–54. https://doi.org/10.3233/BEN-2012-110220

Thompson, C. K., & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology*, *19*(10–11), 1021–1036. https://doi.org/10.1080/02687030544000227

Thompson, C. K., & Shapiro, L. P. (1995). Analysis of verbs and verb-argument atructure: A method for quantification of aphasic language production. *Clinical Aphasiology*, *23*(January 1995), 121–140.

Tippett, D. C., Hillis, A. E., & Tsapkini, K. (2015). Treatment of Primary Progressive Aphasia. *Current Treatment Options in Neurology*, *17*(8), 1–14. https://doi.org/10.1007/s11940-015-0362-5

Ulatowska, H. K., Freedman-Stern, R., Doyel, A. W., Macaluso-Haynes, S., & North, A. J. (1983). Production of narrative discourse in aphasia. *Brain and Language*, *19*, 317–334. https://doi.org/10.1016/0093-934X(83)90074-3

Ulatowska, H. K., Olness, G. S., Wertz, R. T., Samson, A. M., Keebler, M. W., & Goins, K. E. (2003). Relationship between discourse and Western Aphasia Battery performance in African Americans with aphasia. *Aphasiology*, *17*(5), 511–521. https://doi.org/10.1080/0268703034400102

Ulatowska, H. K., Weiss-Doyell, A., Freedman-Stern, R., & Macaluso-Haynes, S. (1981). Production of Narrative and Procedural Discourse in Aphasia. In *Brain and Language* (Vol. 19, pp. 317–334). http://aphasiology.pitt.edu/684/1/10-02.pdf

Van Riper, C. (1947). *Speech Correction: Principles and Methods*. Englewood Cliffs: Prentice-Hall.

Vander Woude, A. (2017). *Connected Language in Primary Progressive Aphasia: Testing the Utility of Linguistic Measures in Differentially Diagnosing PPA and its*

Variants. University of Maryland, College Park.

Vogel, A. P., Poole, M. L., Pemberton, H., Caverlé, M. W. J., Boonstra, F. M. C., Low, E., Darby, D., & Brodtmann, A. (2017). Motor speech signature of behavioral variant frontotemporal dementia: Refining the phenotype. *Neurology*, *89*(8), 837–844. https://doi.org/10.1212/WNL.0000000000004248

Wambaugh, J. L., & Mauszycki, S. C. (2010). Sound production treatment: Application with severe apraxia of speech. *Aphasiology*, *24*(6–8), 814–825. https://doi.org/10.1080/02687030903422494

Wambaugh, J. L., Nessler, C., Cameron, R., & Mauszycki, S. C. (2012). Acquired apraxia of speech: the effects of repeated practice and rate/rhythm control treatments on sound production accuracy. *American Journal of Speech-Language Pathology*, *21*(Select Papers from the 41st Clinical Aphasiology Conference), S5–S27.

Wambaugh, J. L., West, J. E., & Doyle, P. J. (1998). Treatment for apraxia of speech: effects of targeting sound groups. In *Aphasiology* (Vol. 12, Issue 7/8, pp. 731–743).

Weintraub, S., Mesulam, M. M., Wieneke, C., Rademaker, A., Rogalski, E. J., & Thompson, C. K. (2009). The northwestern anagram test: Measuring sentence production in primary progressive aphasia. *American Journal of Alzheimer's Disease and Other Dementias*, *24*(5), 408–416. https://doi.org/10.1177/1533317509343104

Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management*. Grune and Stratton.

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., Miller, B. L., & Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069–2088. https://doi.org/10.1093/brain/awq129

Youmans, G., Holland, A., Muñoz, M. L., & Bourgeois, M. (2005). Script training and automaticity in two individuals with aphasia. *Aphasiology*, *19*(3–5), 435–450. https://doi.org/10.1080/02687030444000877

Youmans, G., Youmans, S. R., & Hancock, A. B. (2011). Script training treatment for adults with apraxia of speech. *American Journal of Speech-Language Pathology*, *20*(1), 23–37. https://doi.org/10.1044/1058-0360(2010/09-0085)